

|CURIEUX|

ACADEMIC JOURNAL

May Issue

Part 2 Issue 50

Editing Staff

Chief-In-Editor

Caroline Xue

George Hasnah

Chief of Operations

Anshal Vyas

Assisting Editors

Olivia Li

Luzia Thomas

Madalyn Ramirez

Shefali Awasthi

Soumya Rai

Sarah Strick

Table Of Contents

Page 6: Dionysian Worship in The Bacchae and Ancient Art By Saanvi Seela

Page 13: Optimizing Stock Market Predictions and Automated Trading with Machine Learning By Ariaman Ladha

Page 27: Behavioral Biases of Consumers in the U.S. During the COVID-19 By Ishana Katriar

Page 44: Decoding Moche Civilisation: The Significance of Stirrup Spout Bottles By Avani Radheshyam

Page 64: Painting a False Picture: Switzerland's Economic and Military Activity with the Nazis & Allies Overlooking their Declared Neutrality During World War II, 1938-1945 By Wanli Zhang

Page 70: AI Maturity in Fortune 500 Firms: Lessons From Cluster Analysis By Abhiram Jetty and Shreyas Chittimalla

Page 82: Structural Analysis of Virulence Factors Between S. aureus Strains MSSA476 and MRSA252 By Linda Yu

Page 92: The Application of the SIR Framework to Analyze COVID-19 Transmission Patterns By Nipun Saxena

Page 110: Rethinking Education for Foster Children: Addressing Struggles Through Comprehensive Educational Strategies By Jolie Zhou

Page 132: Exploring Wax Blends with UV Stabilizers and Antioxidants for Enhanced Solar Panel Applications By Kirti Ponnaganti

Page 144: Mental Health Trends in Just Dance By Manya Jain

Page 148: AI-Powered Tax Insights and Predictions for a Multilingual Audience By Malcolm L. Gao and Chenge Zhang

Page 158: Assessing the Impact of the Hungry Healthy Free Kids Act on Child Health and Food Security By Justin Chen

Page 174: Ketamine's Rapid-Acting Antidepressant Effects are Mediated in Part Through Anti-Inflammatory Mechanisms By Roger Li

Page 188: Market Viability of Electric Aircraft: Overcoming Challenges and Unlocking Potential By Henry M Kong

Page 206: Economic Concerns Perpetuating the Occurrence of Type II Diabetes in AI/AN Communities in the United States By Aabha Vadapalli

Page 212: Adaptive Strategies for Sustaining Corn Yields in Iowa: Assessing the Impact of Climate Change on Precipitation and Temperature Variability By Jiahao Wen

Page 231: Examining the Effect of Narratives on the Persistent Appeal of the Gold Standard By Michael Raso

Page 251: How do Social Norms and Peer Effects Affect Household Debt Accumulation and Financial Stability? A review study By Shraddha Balaji

Page 262: Assessing the Efficiency of Data Collection: A Case Study for the Royal Bengal Tiger By Virenpratap Singh Bhaika

Page 270: A Regression Study of EV in Developed Countries By Yijie Fu

Page 284: Air Pollution and Air Quality Index(AQI) Trends and Predictions Using AI Model By Rishan Pawar

Page 293: Improved Food Safety in Manhattan Restaurants with Data Science backed Guide to Catch Pest Violations By Rohin Vig

Page 309: The Imperative of Ethical Leadership: Lessons from the Mahabharata on How King Dhritarashtra's Bias, Poor Counsel, and Selfish Governance Led to Political Instability, Economic Decline, and the Suffering of His People – Parallels with Modern Political and Business Failures By Varun Ganesh

Page 323: Long-Term Effects of Thymectomies and Congenital Heart Disease on Immune Function and T-Cell Composition By Aarush Anand

Page 333: What Optimisations Can Be Made to RTGS to Improve Their Functionality In Their Current Applications and for Alternative Uses? By Chengxiu Li

Page 340: The Effectiveness of Cognitive Behavioral Therapy in Reducing Symptoms Like Hallucination of Young Adults Diagnosed With Schizophrenia During the Prodromal Stage By Roland Yang

Page 348: What Makes the Most Expensive Contemporary Paintings in the World so Valuable and How Might Their Value Change? By Andrew Zheng

Page 356: Protein Evolution in Humans By Abhay Goel

Page 363: Battling Parkinson's Disease: Cause, Affection, Symptoms, Treatments By Arya Devanath

Page 383: The Journey to Alzheimer's Disease Diagnosis and Beyond Through Different Perspectives By Deetya Potakamuri

Page 396: Causes of Non-Verbal Learning Disorder and the Need for Inclusion in the DSM-5-TR By Advaita Vemuri and Zahara Khan

Page 414: Water Pollution By Dev Shah, Jivin Mahimkar, Siddhanth Dadhich, Veer Furia, and Hriday Shah

Page 423: Ecclesiastical and Classical Latin: A Comprehensive Study of Linguistic Evolution and Theological Implications By Suyu Fang

Page 429: Emerging Middle Powers: The Geopolitical Strategies of Türkiye and Azerbaijan Amidst the Russo-Ukrainian War Emil Javanshir

Page 442: Re-Interpreting the Sent Down Youth: Heidegger, Arendt, and the Vicious Cycle of Worldlessness in the Cultural Revolution By En (Chloe) Ke

Page 450: Cybersecurity Reimagined: Data-Driven Defenses for the Modern Digital Landscape By Ahil Thendral

Page 468: Comparative Analysis of Koch Snowflake and Sierpiński Triangle Fractals: Applications in Optimized Surface Area and Energy Storage of Lithium-ion Battery Electrodes By Ishan Puri

Page 479: The Ethical Implications of Gene Editing Technologies Like CRISPR in Human Medicine By Mingyu Wang

Page 382: Medical Assistance in Dying for Dementia Patients By Aurora Harris

Page 489: Rescuing Wall Street While Overlooking Main Street? A Comparative Analysis of the 2008 Financial Crisis Response By Krishna Rathuryan

Page 496: Global Trade and Domestic Disparities: How U.S. Policy Shapes South Korean Inequality By Jason Chae

Page 508: Sustainable System for Removing Toxic Skid Marks from F1 Racing on Singapore Streets through Waste Upcycling By Benjamin Frederick Oh

Page 522: Independent Research Project on Affect Theory in *Kathak* Performance By Aditi Dhupkar

Page 536: The Hidden Crisis: Innovative Strategies for Climate-Induced Distress in Low Socioeconomic Populations By Riya Kansagra

Page 549: History, Biology, and Advances in the Treatment of Alzheimer's Disease By Ethan Lao

Page 560: The Influence of Color on the Human Mind: An exploration on the Cognitive and Emotional Responses By Tehreem Arshad

Page 571: Can Sanskrit Be Used for Reducing the Energy in LLM Inference Serving By Reducing the Amount of Computation Required By the LLM? By Aanya Bulusu

Page 576: Entering a paradigm-breaking era: How far will AI transform the Economy? By Shaoyang Zhang

Page 581: Fake News Detection Using Neural Network Machine Learning Model By Agrim Thakur

Page 587: Machine Learning to Predict Sustainable Energy Usage Outcomes By James Zhang

Dionysian Worship in The Bacchae and Ancient Art By Saanvi Seela

Introduction

Throughout history, artistic representations of myth have played a key role in conveying complex cultural and philosophical ideas. Among these, Dionysian imagery—depictions of the Greek god Dionysus, his followers, and the rituals associated with him—stands out for its duality: joy and chaos, ecstasy and destruction. Dionysus, god of wine, fertility, and religious ecstasy, was revered not only for his ability to inspire happiness but also for the unnerving powers he unleashed when his divine authority was questioned. The art and literature that portray him are vivid testaments to these competing forces, capturing the tension between liberation and disorder that lies at the heart of this myth.

One of the most profound literary explorations of Dionysus' power is Euripides' *The Bacchae*. Written in the 5th century BCE, the play explores the conflict between human rationality and the unconquerable forces of nature, represented by the god himself. The story centers on Dionysus's return to Thebes, where his mortal cousin Pentheus refuses to recognize his divinity. As Dionysus takes revenge by driving the women of the city into a frenzy, the boundaries between civilization and chaos blur, leading to a climax as terrifying as it is inevitable: Pentheus being killed by his own mother (Euripides, *The Bacchae* 1120-1135). This paper examines how ancient Greek and Roman artworks reflect the themes of Dionysian duality, madness, and transformation found in *The Bacchae*, exploring how artistic portrayals of Dionysian rituals and figures embody the complex, paradoxical themes of the play. Through a comparative analysis of artworks ranging from pottery and mosaics to frescoes and sarcophagi, we will explore how artists across centuries captured the tension between order and chaos embodied by the god, celebrating both the ecstasy of Dionysian liberation and the destruction left in its wake.

Background on Dionysian Worship

To understand Dionysian imagery, we must first look at the god himself and the nature of his worship. Dionysus, also known as Bacchus and Liber, was not just the god of wine; he was the embodiment of a profound, almost elemental force within the Greco-Roman world. Wine represented more than a mere indulgence—it was a symbol of life's dualities. Just as wine could bring joy and release, it also had the potential to bring chaos, violence, and the dissolution of social order. Dionysus, in his essence, embodied these contradictions (Otto 75; Burkert 223).

Central to Dionysian worship was the practice of ecstatic ritual. Followers of Dionysus, particularly his female devotees known as maenads, would abandon societal norms and retreat into the wild, where they would engage in frantic dances and rites, often under the influence of wine. Music and rhythm were integral to these ceremonies, providing a means of accessing altered states of consciousness. These altered states reflected the liminal nature of Dionysus, blurring the boundaries between human and divine, civilized and wild.

At the heart of Dionysian worship is the concept of transformation, both literal and metaphorical. The maenads, for instance, would enter a trance-like state, breaking free of their societal identities and becoming wild and untamed. This transformation speaks to the broader tension in Dionysian worship: the need to release oneself from the constraints of order, only to confront the dangers that lie beyond.

This duality—joy and chaos—is key to understanding *The Bacchae*. Dionysus, as portrayed by Euripides, is not simply a god of liberation. He is a god of reckoning, one who punishes those who fail to acknowledge the chaotic forces he represents. By examining the play itself, we see how these themes of ecstasy and destruction are mirrored in the fates of those who refuse to recognize the god's power (Segal 58).

***The Bacchae* and Its Themes**

At its core, *The Bacchae* is a play about boundaries—between order and chaos, rationality and madness, the human and the divine. Euripides presents Dionysus as a god who demands recognition, not only of his divine status but also of the primal forces he embodies. When Pentheus, king of Thebes, denies Dionysus' divinity, he sets off a chain of events that culminates in his own physically violent yet also symbolic demise.

The central conflict of the play arises from Pentheus' refusal to allow Dionysian worship in Thebes. Pentheus's rational and authoritarian nature represents the rigid structures of law and order. He cannot comprehend the ecstatic frenzy that Dionysus brings, nor does he understand why the women of Thebes—who have been drawn into the mountains to worship the god in wild rituals—are so entranced by him. Pentheus' attempt to reassert control over this chaos is doomed from the start, as Dionysus, in disguise, easily outwits him (Dodds 92).

The climax of *The Bacchae* comes when Dionysus convinces Pentheus to disguise himself as a woman and spy on the Bacchic rites. What follows is not only a personal transformation but a profound psychological unraveling. Pentheus, caught between his desire to control the maenads and his voyeuristic fascination with their freedom, is ultimately torn apart by them—led by his own mother, Agave, who in her madness believes she is killing a wild animal, exclaiming in her delusion, “We caught the lion cub without a trap!” (*The Bacchae*, 1168–1170). Her words here capture the full extent of her tragic frenzy and her inability to recognize her own son (Foley 103).

There are several deeply intertwined themes in this play: the fragility of human order in the face of divine chaos, the power of ecstasy to dissolve boundaries, and the peril of denying one's own nature. Dionysus, both as a character and as a force in the world, exemplifies this tension. He is both liberator and destroyer, rewarding his followers with joy and punishing those who resist with madness and death. This exploration of the thin line between civilization and chaos is not only central to *The Bacchae* but also to the visual depictions of Dionysian worship in art. In these images, we see the same transformative power and delicate balance—whether in the hectic dance of the maenads or the moment of Pentheus' violent undoing.

Symposium: Dionysian Joy and Restraint in Greek Art

The Greek symposium, a social and cultural gathering for philosophical discussion and revelry, served as a space where the rituals of Dionysus intersected with artistic expression. Wine was central to the symposium, both as a symbol of Dionysian joy and as a reminder of the need for balance and restraint. Mythological imagery often adorned the vessels used in these gatherings, reflecting the dualities of life and celebrating the god who presided over both joy and chaos.

A key artifact that exemplifies this is the bell krater, a type of large vessel used for mixing wine and water. One such bell krater from around 400 BCE is decorated with a scene filled with Dionysian imagery: maenads and satyrs dance in wild abandon, holding thyrsi and drinking vessels, their expressions and movements suggestive of a world outside the bounds of ordinary life (Boardman 157–160). The thyrsus, a recurring motif, represents both fertility and violence—a weapon of transformation. It is held by Dionysus and his followers not just as a symbol of life's pleasures but also as a powerful reminder of the potential for chaos when order is disrupted. The maenads, with their frenzied movements, embody the ecstatic release that comes from surrendering to the god's power, their dance echoing the chaos that erupts in *The Bacchae* when Pentheus attempts to suppress the Dionysian rites. The bell krater's imagery serves as a visual parallel to the themes of *The Bacchae*, visually illustrating how Dionysus' rites encourage both ecstasy and danger (Shapiro 98).

A similar artifact is a red-figure cup in the Kimbell Art Museum, attributed to Douris and dating to c. 480 BCE. The exterior depicts the violent climax of *The Bacchae*: Pentheus' dismemberment at the hands of frenzied maenads. His fragmented body is shown amidst dynamic, chaotic movements. The interior, in contrast, portrays a serene maenad holding a thyrsus and dancing, suggesting the celebratory aspect of Dionysian worship (Boardman 172). This juxtaposition reflects the duality inherent in Dionysian rites: the ecstasy of joy and the peril of chaos.

These vessels mirror the themes of *The Bacchae*. The maenads' dances suggest the liberating ecstasy of surrendering to Dionysus, yet their frenzied movements warn of the chaos that ensues when this ecstasy goes unchecked. Pentheus' refusal to honor Dionysus ultimately leads to his unraveling. These artifacts illustrate how Dionysian imagery served both as tools for revelry and as reminders of the delicate balance between joy and destruction (Lissarrague 92).

House Decorations: Immersion in Dionysian Rites

Dionysian imagery also played a central role in Roman domestic spaces, adorning walls and floors in mosaics and frescoes that reflected the god's themes of transformation, joy, and chaos. These decorations immersed viewers in Bacchic rituals, serving both as celebrations of Dionysus and as reminders of his dual nature.

The frescoes of the Villa of Mysteries in Pompeii provide one of the most enigmatic and evocative representations of Bacchic rituals in Roman art. Painted in the 1st century CE, these frescoes adorn the walls of a room thought to have been used for the initiation of individuals into the cult of Dionysus. The imagery in these frescoes is rich with symbolic references to the rites

and mysteries associated with Dionysian worship, and they offer a glimpse into the deeply spiritual and transformative nature of these rituals (Elsner 112; De Caro 45–47).

The frescoes depict a series of figures engaged in what appears to be a Bacchic initiation ceremony. Central to the imagery is Dionysus himself, reclining languidly while his followers, including maenads and satyrs, participate in rituals of music, dance, and intoxication. The figures are rendered in flowing forms that convey a sense of ecstatic movement, their gestures and expressions suggesting both reverence and rapture. The rich, deep colors of the frescoes, particularly the reds and golds, evoke the intensity and fervor of the Bacchic rites, imbuing the scene with a sense of mystery and sacredness.

The ritualistic elements of the frescoes are deeply connected to the themes of *The Bacchae*. The initiation into the mysteries of Dionysus often involved a symbolic death and rebirth, echoing the theme of transformation that runs throughout Euripides' play. The figures in the frescoes, particularly the women, seem to be undergoing a similar transformation, their participation in the rites marking their transition from the ordinary world into the realm of Dionysian ecstasy and chaos. The frescoes also highlight the tension between order and chaos that is central to Dionysian worship, visually resonating with *The Bacchae*'s exploration of identity and transformation (Sauer 76).

Similarly, the *Death of Pentheus* mosaic vividly captures the darker side of Dionysian worship. Found in the southern triclinium of the House of the Vettii in Pompeii, this mosaic presents a visceral and detailed depiction of Pentheus' demise at the hands of the Bacchic women, including his mother, Agave. Unlike the more symbolic and stylized approach of earlier Greek art, this Roman piece offers a strikingly literal interpretation, showing Pentheus' body in mid-dismemberment, with the maenads, wild-eyed and driven by Dionysian madness, tearing him apart (Ling 203; Zanker and Ewald 190). The mosaic's composition reflects the chaotic energy of the scene. The maenads are depicted in dynamic poses, their limbs and clothing swirling around them as they engage in the violent act. Pentheus, meanwhile, is caught in the center of this chaos, his body contorted and fragmented, a visual symbol of the disintegration of order under the force of Dionysian ecstasy. The use of bright, contrasting colors further heightens the emotional intensity of the scene, drawing the viewer into the horror and frenzy that consumes Pentheus.

This brutal realism reflects the Roman penchant for narrative and directness in their art, but it also serves as a potent visual metaphor for the core themes of *The Bacchae*: the destruction of order by the forces of chaos, the fragility of human control in the face of divine power, and the inevitable consequences of denying the gods their due. The mosaic captures not only the physical violence of Pentheus' death but also the symbolic destruction of his authority, highlighting the devastating power of Dionysus and his followers (Ling 223; Zanker and Ewald 190). In this way, the mosaic acts as a powerful visual counterpart to Euripides' tragedy, bringing the themes of the play into sharp, literal focus.

Sarcophagi: Dionysus and the Cycle of Life and Death

Roman sarcophagi frequently depicted scenes from Dionysian rites, illustrating the god's association with transformation and the inevitability of death. These carvings linked the themes of Bacchic revelry to the afterlife, suggesting that surrender to Dionysus offered a form of continuity beyond mortal existence.

A third-century CE sarcophagus, adorned with Bacchic imagery, encapsulates the themes of Dionysian chaos and transformation. Unlike the dynamic energy of the maenad reliefs and mosaics, the figures on this sarcophagus appear more static and serene, yet the themes are readily apparent. Carved in high relief, the scenes show maenads and satyrs in various stages of celebration, their postures suggesting a quiet, eternal revelry. Some of these figures are frozen in ecstatic poses, holding thyrsi or wine vessels, symbolizing the themes of indulgence and freedom (Koch 105; Borg 87).

While Greek art frequently represents Dionysian rites as temporary releases from the structured life of the polis, Roman sarcophagi suggest a more enduring connection between life and the divine. The theme of transformation here is closely tied to the sarcophagus's primary purpose: housing the dead. Through Bacchic imagery, these sarcophagi imply that death, like Dionysian worship, is a transition—a passage from one state of existence to another. In *The Bacchae*, transformation is both the reward and the punishment bestowed by Dionysus, as he breaks down the walls of self-control and rationality in Thebes. This sarcophagus reflects the broader scope of Dionysian worship, linking it to the cycle of life and death, symbolizing that surrender to Dionysian ecstasy might offer solace or transformation beyond the mortal realm (Sichtermann 120).

In the context of *The Bacchae*, the use of Bacchic imagery on sarcophagi also hints at the immortality of divine ecstasy. Just as Pentheus' death symbolizes the power of Dionysian madness over human order, the Bacchic scenes on sarcophagi suggest that submission to Dionysus offers a kind of release or continuity beyond death. This visual emphasis on eternal revelry aligns with the thematic elements of *The Bacchae*, where Dionysus is portrayed as a force that is both boundless and enduring, transcending mortal limitations.

The sarcophagi's quiet reverence reflects Dionysus' role as a guide through life and death. Just as Pentheus' destruction in *The Bacchae* symbolizes a forced transformation through divine madness, these sarcophagi suggest that surrender to Dionysian ecstasy leads to a different kind of transformation—one that bridges the mortal and divine (Sichtermann 120). By linking Bacchic rites to the afterlife, these sarcophagi portray Dionysus as both a liberator and a harbinger of inevitable change.

Comparative Analysis

Depictions of Dionysian worship reveal the god's complex and unsettling duality: he embodies both ecstatic joy and a force of divine reckoning. In *The Bacchae*, this paradox is central, with Dionysus portrayed as a god who offers his followers blissful freedom through music, dance, and wine, dissolving the boundaries of social order and releasing them into states of euphoric abandon. This aspect of Dionysian worship is often represented in art through the

lively movements of the maenads, as seen on the Bell Krater, where their dance and flowing garments capture the liberating, intoxicating power of the god (Boardman 157). Such images invite viewers into the ecstatic celebration of Dionysus, emphasizing the joy and unity he brings to his followers.

Yet Dionysus's energy carries a darker, more threatening edge. Beneath the joyful façade is a god who demands reverence and, when denied, responds with devastating retribution. This aspect is vividly depicted in the Death of Pentheus mosaic, where Dionysus's power drives the maenads into a frenzy that ultimately leads to Pentheus's brutal dismemberment, a visual reflection of divine wrath upon those who challenge his authority (Ling 203; Zanker and Ewald 190). The same god who brings joy and liberation also unleashes chaos and destruction upon those who resist him, as seen in the Villa of Mysteries frescoes, where the transformative Bacchic rites imply both a symbolic rebirth and the underlying potential for psychological unraveling (Elsner 112).

This dual nature of Dionysus emphasizes the balance between joy and terror. The god's ecstatic power brings freedom but also holds the potential for uncontrollable violence. As these artworks suggest, Dionysian worship is a precarious dance on the edge, where the line between blissful ecstasy and consuming madness is thin. The Bell Krater, the Villa of Mysteries frescoes, and the Death of Pentheus mosaic collectively illustrate this duality, capturing the allure of Dionysus's gifts and the future that awaits those who fail to respect his transformative, often unpredictable nature.

Works Cited

- Boardman, John. *Greek Vases: A Handbook*. Thames & Hudson, 2001.
- Burkert, Walter. *Greek Religion*. Harvard University Press, 1985.
- De Caro, Stefano. *The Frescoes of the Villa of Mysteries*. Getty Publications, 1995.
- Dodds, E. R. *Euripides' Bacchae*. Oxford University Press, 1960.
- Elsner, Jas. *Roman Eyes: Visuality and Subjectivity in Art and Text*. Princeton University Press, 2007.
- Euripides. *The Bacchae*. Translated by Paul Woodruff, Hackett Publishing, 1998.
- Foley, Helene P. *Ritual Irony: Poetry and Sacrifice in Euripides*. Cornell University Press, 1985.
- Koch, Guntram. *Roman Sarcophagi*. Springer, 1993.
- Lissarrague, François. *The Aesthetics of the Greek Symposium*. Princeton University Press, 1990.
- Otto, Walter F. *Dionysus: Myth and Cult*. Indiana University Press, 1965.
- Segal, Charles. *Dionysiac Poetics and Euripides' Bacchae*. Princeton University Press, 1997.
- Sichtermann, Hellmut. *Roman Sarcophagi in Context*. Getty Publications, 2003.
- Zanker, Paul, and Björn Ewald. *Living with Myths: The Imagery of Roman Sarcophagi*. Oxford University Press, 2012.

Optimizing Stock Market Predictions and Automated Trading with Machine Learning

By Ariaman Ladha

Abstract

This study explores the development and implementation of an artificial intelligence trading bot designed to predict stock prices and automate trading decisions, leveraging machine learning techniques. Utilizing historical stock data from the US stock market, the research employs a range of models, including linear regression, decision tree regression, random forest regression, and neural networks. Each model's unique abilities are combined to enhance predictive accuracy, with predictions integrated using a multiplicative weight update method that dynamically adjusts based on performance. The trading bot is evaluated through conservative and aggressive strategies, each tailored to different risk profiles and market conditions. The conservative strategy returned an average profit of 448%, returning nearly six times the initial investment. Similarly, the aggressive model returned an average profit of 415%, with a more modest yet significant five-time return. This comprehensive approach not only aims to improve the accuracy of stock price predictions but also provides a practical tool for real-time trade automation, advancing the application of AI in financial markets and offering valuable insights for investors seeking to optimize their trading portfolios.

Introduction



The application of artificial intelligence in financial markets has significantly evolved over the past few decades, offering new methods for analysing data, predicting market trends,

and automating trading processes. AI-driven trading systems, often referred to as trading bots, utilise sophisticated machine learning algorithms to interpret vast amounts of financial data, identify patterns, and execute trades with minimal human intervention. This integration of technology has the potential to revolutionise trading by improving accuracy, eliminating emotional biases, and enhancing overall efficiency. Although AI isn't an exact predictor of stock prices as of yet, many studies have investigated different methods of stock price prediction. These studies indicate promise in the possibility of developing an accurate machine-learning trading bot.

Predicting stock and stock price index movement using Trend 4 Deterministic Data Preparation and machine learning techniques (Patel et al.,) is an example of this research. This paper explores trend-deterministic data preparation and machine-learning techniques for predicting stock prices. It aims to enhance the accuracy of stock price predictions by meticulously preparing data to capture underlying trends more effectively. They employed various machine learning algorithms, including neural networks and support vector machines, to analyse historical stock data and predict future price movements. The study demonstrated that incorporating trend deterministic data preparation significantly improved the predictive performance of these models. However, the authors identified several challenges, such as the high computational complexity involved in processing large datasets and the necessity for substantial computational resources to train complex models. Additionally, they emphasised the critical role of feature selection and data preprocessing in enhancing model accuracy. Despite these challenges, the study underscored the potential of machine learning techniques in stock price prediction, highlighting the importance of efficient data handling and advanced preprocessing methods to achieve optimal results.

Prediction model for stock price trend based on recurrent neural networks (Zhao et al., 2020) investigated the use of recurrent neural networks (RNNs) for predicting stock prices, capitalizing on the sequential nature of financial data. Their study highlighted that recurrent neural networks are particularly effective in capturing temporal dependencies within stock price movements, leading to enhanced prediction accuracy over traditional models that do not consider time-series characteristics. Despite the advantages, the authors pointed out significant challenges, such as the high computational power required for training and the propensity for overfitting with complex RNN structures. They suggested that combining RNNs with other techniques like regularization and ensemble methods could improve model robustness and accuracy. This research contributes valuable insights into the potential of advanced neural network architectures in financial forecasting and underscores the need for continued innovation in AI-driven financial analysis.

Stock Market Forecasting Using Machine Learning Algorithms (Shen et al., n.d.) is a study on stock market forecasting using machine learning algorithms, focusing on the practical application of these techniques in predicting stock prices. The authors compared the performance of several machine learning algorithms, including support vector machines, k-nearest neighbours, and artificial neural networks, in forecasting stock market trends. Their study

highlighted that machine learning algorithms could significantly enhance prediction accuracy compared to traditional statistical methods. One of the key contributions of this research was its detailed analysis of the algorithmic differences and their implications for stock market forecasting. The authors demonstrated that while artificial neural networks excelled in capturing the complex, non-linear relationships within the data, support vector machines and k-nearest neighbours provided valuable insights into different market patterns. However, they also noted the inherent challenges associated with each algorithm, such as the susceptibility of neural networks to overfitting and the sensitivity of k-nearest neighbours to data scaling and noise. This study provided a comprehensive comparison of different machine learning approaches, offering valuable insights into their practical applications and limitations in the context of stock market prediction.

Building on the insights and methodologies from these previous studies, this would like to advance the field of algorithmic trading by not only predicting stock prices but also automating the trading process in various levels of aggression based on these predictions. By developing an integrated trading bot, this study seeks to further previous research by providing a practical tool that can execute trades autonomously. This approach aims to enhance the overall efficiency and effectiveness of trading strategies, bridging the gap between theoretical advancements in machine learning and their practical application in stock market trading. The goal is to contribute valuable insights and tools to the field of machine learning-driven financial analysis, pushing the boundaries of what is possible with current technology.

Methodology

Developing an AI-driven trading bot for predicting stock prices and automating trades involves a comprehensive methodology that includes data collection, preprocessing, model selection, training, evaluation, and implementation. This section details each step, focusing on the techniques and algorithms used to achieve accurate and reliable predictions.

Data Collection

The dataset for this project includes all stocks listed in the US stock market. Historical stock data is retrieved using the Yahoo Finance library, which provides easy access to financial data. The primary focus is on the daily market open prices of stocks over a five-year period. Data is collected for five different stocks to ensure a comprehensive dataset of historical open prices.

Data Preprocessing

Preprocessing is crucial to ensure the data being used is suitable for model training. This involves several key steps:

1. **Creating Input-Output Pairs:** The open prices from the previous three days (input) are used to predict the next day's open price (output). This transformation ensures that the model captures temporal dependencies and trends in the stock prices, which are crucial for accurate predictions.

The input features consist of the open prices for three consecutive days, and the output is the open price for the following day.

2. Splitting the Dataset: The dataset is split into training and testing sets. Specifically, 67% of the data is used for training the models, and the remaining 33% is used for testing. This split is important to evaluate the models' performance on unseen data and ensure that they generalize well to new data. This method provides a balanced approach to model training and validation, allowing for a robust evaluation of model performance.

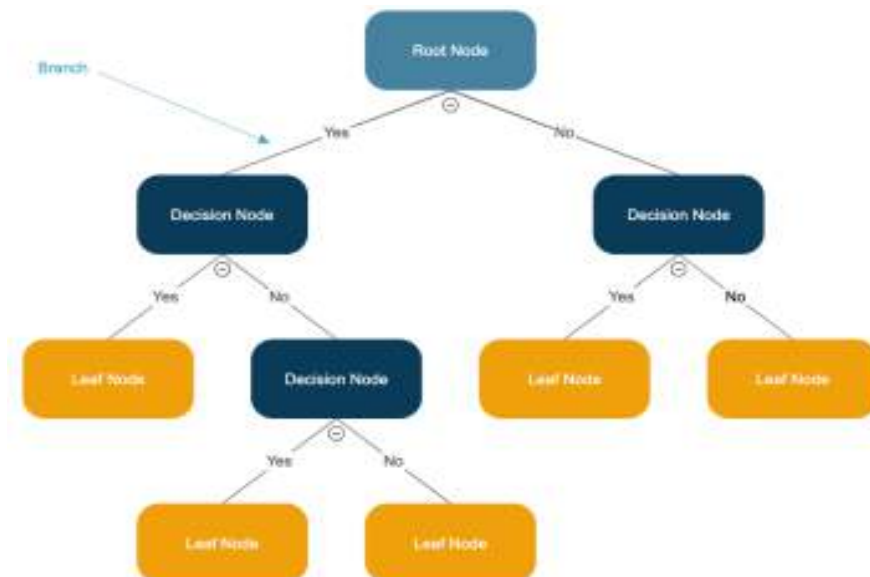
Model Selection and Training

Several machine learning models are employed to predict stock prices, each handling a unique function. The goal is to leverage all of these unique functions and improve overall prediction accuracy, thus maximizing profit. The models used in this project include linear regression, decision tree regression, random forest regression, neural networks, and a model that combines all these.

Linear Regression

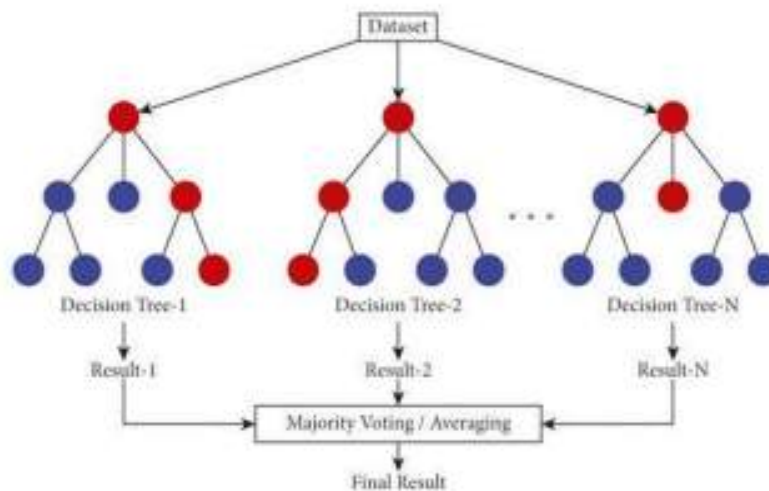
Linear regression is a fundamental statistical technique that models the relationship between a dependent variable and one or more independent variables. It is particularly useful for capturing linear relationships. Linear regression is advantageous due to its simplicity and ease of interpretation, making it an excellent starting point for understanding the relationships in the data. However, its primary limitation is its inability to capture non-linear relationships, which are common in financial data.

Decision Tree Regression



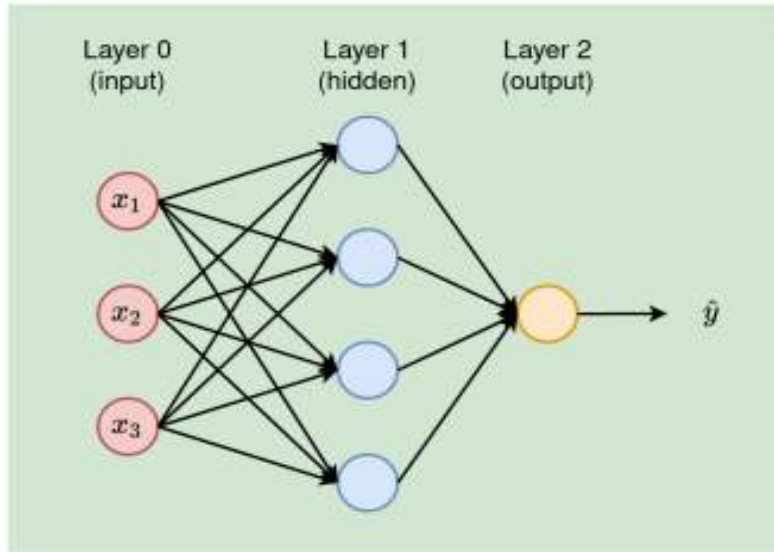
Decision tree regression is a non-parametric supervised learning method used to predict the value of a target variable by learning decision rules from features. It constructs a tree where each internal node represents a feature, each branch represents a decision rule, and each leaf represents a predicted outcome. Decision trees can fit complex patterns by learning local linear regressions, making them capable of approximating functions such as a sine curve with added random errors in the observations. The primary advantages include its simplicity, interpretability, and the ability to handle both numerical and categorical data without requiring feature scaling. However, decision trees are prone to overfitting, especially when the tree is too deep, capturing random errors in the training data. Setting a maximum depth can help mitigate this issue.

Random Forest Regression



Random forest regression is an ensemble method that combines multiple decision trees to improve prediction accuracy and reduce overfitting. By averaging the results of multiple decision trees, random forests mitigate the overfitting problem inherent in single decision trees and provide more robust and accurate predictions. The use of ensemble methods allows the model to capture a broader range of market dynamics and improve generalizability.

Neural Networks



Neural networks are powerful models capable of capturing complex patterns and non-linear relationships in the data. Neural networks excel at capturing intricate and non-linear relationships in large datasets, making them ideal for complex prediction tasks. However, they require substantial computational resources and are prone to overfitting if not properly regularised. A multilayer perceptron (MLP) is used in this project, which consists of multiple hidden layers that process the input data to predict the output.

Model Integration using Multiplicative Weight Update

Predictions from all models are integrated using a multiplicative weight update method to enhance the overall predictive accuracy and capitalize on each model's strengths. This approach adjusts the weights of each model's predictions based on their respective errors, aiming to minimize the overall prediction error.

1. Initialization: Initially, equal weights are assigned to each model. This means that each model's prediction has an equal influence on the final combined prediction. These weights sum up to 1.
2. Prediction Combination: The combined prediction is calculated as a weighted average of the individual model predictions.
3. Error Calculation: The prediction error for each model is calculated as the absolute difference between the predicted and actual values.
4. Weight Adjustment: The weights are adjusted by penalising models with higher errors and rewarding models with lower errors. Specifically, the weight of each model is divided by 2 raised to the power of the error for that model.

5. Normalization: The adjusted weights are then normalized so that their sum remains 1, ensuring that the combined prediction remains balanced.

This process is repeated for every input price in the testing set of the data.

Automating Trades

The final step involves automating trades based on the predictions. The trading bot executes buy and sell orders based on the predicted prices, with the objective of optimizing returns. Two trading strategies are implemented: a conservative strategy and a more aggressive one.

Conservative Trading Strategy

In the conservative strategy, the bot purchases shares equivalent to twenty per cent of the initial investment amount into any given stock when the predicted price is higher than the current price and sells the same number of shares when the predicted price is lower. This approach aims to minimize risk by adjusting the investment size relative to the initial amount, ensuring that the bot trades smaller quantities relative to the portfolio value. This strategy is particularly suitable for stable and gradual returns. By investing a proportion of the initial amount rather than a constant number of shares, this method accounts for variations in stock prices across different companies, providing a more balanced and risk-adjusted investment approach.

Aggressive Trading Strategy

In the aggressive strategy, the bot buys and sells larger quantities of shares based on the predictions. Specifically, the bot invests fifty per cent of the initial amount into shares when the predicted price is higher than the current price and sells the same amount when the predicted price is lower. This strategy seeks higher returns by trading larger volumes, making it suitable for investors who are willing to take on more risk for potentially higher rewards. Similar to the conservative strategy, this proportional investment approach allows for more flexibility and better risk management compared to a fixed number of shares, as it adjusts to the price of the stock, ensuring that the investment size is appropriate relative to the value of the stock.

Evaluation and Validation

The trading bot's performance is evaluated using the mean squared error (MSE). MSE measures the average of the squares of the errors, where the error is the difference between the predicted and actual values. A lower MSE indicates a better fit of the model to the data. The models are validated using a holdout validation method, where the dataset is split into training and testing sets to ensure the robustness of the predictions.

By combining multiple machine learning models and integrating their predictions, this methodology aims to leverage the strengths of each model while mitigating their individual

weaknesses. Linear regression offers simplicity and interpretability; decision trees can fit complex patterns through local approximations, random forests provide robustness and accuracy through ensemble learning, and neural networks excel at capturing intricate patterns in data. The integration of a multiplicative weight update method further enhances predictive accuracy by dynamically adjusting model weights based on their performance.

Automating the trading process based on these predictions allows for real-time trade execution, improving the overall efficiency and effectiveness of the trading strategies. This comprehensive approach not only aims to improve the accuracy of stock price predictions but also provides a practical tool for automating trades in real-time. By advancing beyond prediction to the automation of trading decisions, this study builds on previous research and contributes valuable insights and tools to the field of AI-driven financial analysis.

Results

Conservative Model Performance

The conservative model, designed to minimize risk and prioritize stability, showed mixed results across different stocks. The strategy focuses on gradual, consistent returns by trading smaller quantities and avoiding significant exposure to volatile market movements.

In this model, the code checks whether the price prediction for the next day is higher or lower than the current day. If it is higher, it checks that there is enough balance to purchase five shares and if so, proceeds to purchase them, subtracting from the balance. If it is lower, it sells five shares and adds the sale value to the balance. After all the data has been iterated through and trades have been made for each day, a final amount of money is returned.

Stock Ticker	Initial Amount	Final Amount	Overall Profit/Loss (%)
AAPL	10000	35312.20	+253.12
MSFT	10000	30440.52	+204.40
NVDA	10000	153464.34	+1434.64
GOOG	10000	29721.25	+197.21
AMZN	10000	25124.40	+151.24

AAPL: The conservative model's gain of 253.12% for AAPL can be attributed to Apple's relatively stable and predictable stock behaviour, characterised by consistent growth with occasional volatility spikes. The conservative strategy took advantage of these small fluctuations, steadily accumulating gains without exposing itself to significant risks.

MSFT: The gain of 204.40% in MSFT reflects Microsoft's stable growth trend, which is punctuated by periodic price movements due to product launches and quarterly earnings reports.

The conservative model benefited from these predictable patterns, making small but consistent profits over time.

NVDA: NVDA's impressive gain of 1434.64% demonstrates the model's ability to handle high volatility effectively. Nvidia's stock is known for its rapid price changes, driven by the tech sector's innovations and market demand for GPUs. The conservative model capitalized on these frequent price swings without overexposing itself.

GOOG: The 197.21% gain in GOOG highlights the model's strength in leveraging Google's steady upward trajectory. Google's stock experiences moderate volatility, which the conservative model exploited by making consistent, low-risk trades during minor dips and rallies.

AMZN: The gain of 151.24% for AMZN indicates the model's success in navigating Amazon's highly volatile stock behaviour. Amazon's stock price is influenced by its vast e-commerce operations and fluctuating profit margins. The conservative strategy effectively managed these fluctuations to achieve significant returns.

Overall Performance: The conservative model's average profit/loss across all stocks was +448.12%. This model was effective across various stocks, demonstrating its ability to achieve stable and significant returns.

Aggressive Model Performance

The aggressive model aims for higher returns by trading larger volumes and taking on more risk. Here, a higher initial amount is assigned to have the ability to purchase more shares. This strategy is designed to capitalize on significant price movements, making it suitable for stocks with higher volatility.

In this model, the code checks whether the price prediction for the next day is higher or lower than the current day. If it is higher, it checks that there is enough balance to purchase 1000 shares and if so, proceeds to purchase them, subtracting from the balance. If it is lower, it sells five shares and adds the sale value to the balance. After all the data has been iterated through and trades have been made for each day, a final amount of money is returned.

Stock Ticker	Initial Amount	Final Amount	Overall Profit/Loss (%)
AAPL	10000	15821.70	+58.22
MSFT	10000	21216.12	+112.16
NVDA	10000	164052.72	+1540.53
GOOG	10000	24735.75	+147.36
AMZN	10000	31904.00	+219.03

AAPL: The aggressive model's gain of 58.22% for AAPL is relatively modest due to Apple's stable growth and lower volatility. The aggressive strategy found fewer opportunities to exploit large price movements, resulting in smaller overall gains compared to more volatile stocks.

MSFT: The gain of 112.16% in MSFT reflects the aggressive model's ability to capitalize on Microsoft's consistent yet occasionally volatile performance. The model leveraged larger trades during periods of significant growth, amplifying returns.

NVDA: The substantial gain of 1540.53% for NVDA demonstrates the aggressive model's strength in high-volatility environments. Nvidia's frequent and sharp price changes provided numerous opportunities for large-volume trades, resulting in exceptional profits.

GOOG: The 147.36% gain in GOOG indicates the model's effectiveness in taking advantage of Google's moderate volatility. The aggressive strategy executed larger trades during significant upward movements, leading to substantial returns.

AMZN: The gain of 219.03% for AMZN highlights the model's success in leveraging Amazon's high volatility. The aggressive strategy's larger trades allowed it to exploit significant price movements, achieving considerable gains.

Overall Performance: The aggressive model's average profit/loss across all stocks was +415.46%, demonstrating its effectiveness in achieving substantial profits across all analyzed stocks. By taking on more risk and trading larger volumes, the model capitalized on significant price movements, particularly benefiting from the volatility of stocks like NVDA and AMZN.

Model Comparison

Model	Average Profit/Loss (%)
Conservative	+448.12
Aggressive	+415.46

The comparative analysis reveals that both the conservative and aggressive models achieved substantial average profits. The conservative model, with an average profit/loss of +448.12%, showed its effectiveness in providing stable and significant returns across various stocks. The aggressive model, with an average profit/loss of +415.46%, also demonstrated its ability to achieve substantial profits, particularly in volatile market conditions.

The choice between conservative and aggressive models depends largely on the investor's risk tolerance and market conditions. The conservative model is more suited for stable stocks and risk-averse investors, while the aggressive model is ideal for those willing to take on more risk to achieve higher returns, especially in volatile markets. This study underscores the importance of

tailoring trading strategies to the specific characteristics of the stock and the investor's risk profile, providing valuable insights for AI-driven financial analysis and trading.

Conclusion

This study has demonstrated the transformative potential of artificial intelligence in the realm of stock trading, particularly through the development and deployment of a trading bot capable of both predicting stock prices and executing trades. By leveraging a combination of sophisticated machine learning models, this research has advanced our understanding of how AI can be utilized to enhance trading efficiency, accuracy, and overall profitability.

One of the key contributions of this study is the integration of multiple machine learning models, each with its distinct strengths, to create a robust and reliable prediction system. The use of linear regression, decision tree regression, random forest regression, and neural networks allowed the model to capture a wide array of stock price behaviors. The multiplicative weight update method further refined these predictions, ensuring that the model dynamically adjusted to minimize errors and optimize accuracy.

The conservative model showcased the benefits of a cautious approach to trading. By prioritizing stability and minimizing risk, this model effectively managed to generate consistent returns across various stocks. The model's performance with AAPL and MSFT highlighted its ability to exploit stable growth patterns, while its handling of more volatile stocks like NVDA and AMZN demonstrated its capability to navigate and profit from market fluctuations without significant risk exposure. The overall average profit/loss of +448.12% for the conservative model underscores its efficacy in providing stable, long-term returns.

Conversely, the aggressive model exemplified the potential of high-risk, high-reward trading strategies. This model's approach of trading larger volumes to capitalize on significant price movements proved particularly effective in volatile markets. The substantial gains observed with NVDA and AMZN highlight the aggressive model's strength in leveraging rapid price changes to achieve remarkable profits. However, the model also showed that even in less volatile stocks like AAPL, the ability to execute larger trades can still yield significant returns, albeit to a lesser extent than in more volatile conditions. The average profit/loss of +415.46% for the aggressive model demonstrates its potential for investors willing to embrace higher risks for greater rewards.

The comparative analysis between the conservative and aggressive models provides critical insights into the importance of aligning trading strategies with investor risk profiles and market conditions. While the conservative model is ideally suited for risk-averse investors and stable market environments, the aggressive model offers substantial benefits for those willing to take on more risk in pursuit of higher returns, particularly in volatile markets. This study thus highlights the necessity of personalized trading strategies, tailored to both the specific characteristics of the stock and the investor's risk tolerance.

Moreover, this research goes beyond mere stock price prediction by automating the trading process, showcasing a practical application of AI in real-time trading. The trading bot

developed in this study represents a significant leap forward in trading technology, bridging the gap between theoretical advancements in machine learning and their practical implementation in the financial markets. This automation not only enhances trading efficiency but also reduces the potential for human error and emotional bias, leading to more disciplined and systematic trading practices.

In conclusion, this study underscores the substantial potential of AI-driven trading bots in transforming stock trading practices. By integrating multiple machine learning models and automating the trading process, this research offers a comprehensive and practical tool for investors seeking to navigate the complexities of the stock market with greater precision and efficiency. The insights gained from this study provide a solid foundation for future advancements in AI-driven financial analysis, paving the way for even more sophisticated and effective trading systems.

Possible Extensions to the Study

Incorporating Quarterly Financial Reports

One possible extension to this study is the inclusion of quarterly financial reports in the price prediction algorithm. Financial reports provide comprehensive insights into a company's financial health, including revenue, profits, expenses, and other critical metrics. By integrating this information, the model could better understand the underlying factors driving stock prices and make more informed predictions.

For instance, incorporating earnings per share (EPS), revenue growth, and profit margins could help the model predict how a stock might react to financial disclosures. This integration could be achieved by combining time-series data from stock prices with quarterly financial data, potentially improving the model's accuracy and robustness. The enhanced model would be able to account for both historical price movements and the latest financial performance, providing a more holistic view of the stock's potential future behavior.

Including Financial News in the Price Prediction Algorithm

Another extension could involve incorporating financial news into the price prediction algorithm. News articles, press releases, and social media sentiment can significantly impact stock prices, as they often contain information about market trends, company performance, and investor sentiment.

To implement this, natural language processing (NLP) techniques could be used to analyse the sentiment and relevance of news articles related to the stocks being traded. By scoring news articles based on their sentiment (positive, negative, or neutral) and relevance, the model could adjust its predictions accordingly. This approach would enable the trading bot to react to real-time information and anticipate market movements driven by news events, potentially increasing its responsiveness and accuracy in volatile markets.

Integrating financial news with historical price data and quarterly financial reports could create a comprehensive prediction model that leverages multiple data sources. This enhanced

model would be better equipped to navigate the complexities of the stock market, providing more accurate predictions and improving trading decisions.

These extensions would further advance the capabilities of the AI-driven trading bot, making it more robust and versatile in handling a wide range of market conditions and data sources. By continually refining and expanding the model's inputs, the trading bot could achieve even greater accuracy and effectiveness, offering valuable insights and tools for investors in the ever-evolving financial markets.

Works Cited

- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259–268.
<https://doi.org/10.1016/j.eswa.2014.07.040>
- Shen, S., Jiang, H., & Zhang, T. (n.d.). Stock Market Forecasting Using Machine Learning Algorithms.
<https://cs229.stanford.edu/proj2012/ShenJiangZhang-StockMarketForecastingusingMachineLearningAlgorithms.pdf>
- Zhao, J., Zeng, D., Liang, S., Kang, H., & Liu, Q. (2020). Prediction model for stock price trend based on recurrent neural network. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-020-02057-0>
- Colvin, G., Hemann, G., & Kalouche, S. (n.d.). Stock Forecasting using Machine Learning. Retrieved August 1, 2024, from <https://web.stanford.edu/~kalouche/docs/AI.pdf>

Behavioral Biases of Consumers in the U.S. During the COVID-19 By Ishana Katriar

Pandemic Abstract

This paper examines the impact of behavioral biases on consumer behavior during the COVID-19 pandemic, focusing on herd mentality, loss aversion, mental accounting, overconfidence bias, and the availability heuristic. The study aims to understand which biases were prominent in consumer behavior during the pandemic and how they manifested. The study reveals that herd mentality intensifies due to increased social media usage, driving trends like panic buying and the popularity of specific products. Loss aversion influenced cautious financial behavior during this stressful time, as individuals prioritized savings over spending in response to perceived inflationary pressures. Overconfidence bias was apparent in cryptocurrency investments, where investors underestimated risks despite volatile conditions. The availability heuristic skewed consumer perceptions of necessity and value based on the prominence of information. On the other hand, the less evident bias was Mental accounting; as immediate financial needs led to a blending of different funds, it was contrary to the typical behavior of treating stimulus payments as separate from regular income. These insights can contribute to a deeper understanding of how behavioral biases emerge and influence consumer behavior during crises, offering implications for managing economic behavior in future hardships

1. Introduction

The COVID-19 pandemic (COronaVirus Disease), which began in late 2019, rapidly escalated into a global crisis, disrupting nearly every aspect of daily life. The World Health Organisation (WHO) declared the virus a global pandemic on March 11, 2020 and issued lockdowns (World Health Organization, 2020). As countries implemented social distancing measures to mitigate the spread of the virus, consumer behavior underwent significant transformations. From panic buying to changes in spending priorities, the pandemic exposed underlying psychological mechanisms that drive economic decision-making (Goolsbee and Syverson, 2021).

Behavioral economics provides a framework for understanding these changes, particularly through behavioral biases. Behavioral economics, a field of economics derived in the 1980s due to research by Daniel Kahneman, Amos Tversky, and Richard Thaler, combines economics and psychology to understand how people behave in the real world instead of rationally (Kahneman and Tversky, 1979; Thaler, 1980). This branch of economics has specific biases that allow individuals to deviate from the standard economic models that assume rational decision-making. Instead, it highlights the mental shortcuts and heuristics people use in uncertain situations (Tversky and Kahneman, 1974). In a rational economic model, consumers are expected to make decisions that maximize their utility based on all available information. However, in practice, various biases can lead to irrational behaviors, especially during crises and periods of high stress (Thaler, 1985).

In addition to the psychological effects of the pandemic, economic behavior was also

influenced by significant fiscal and monetary policies. The U.S. government implemented several interventions, including the Coronavirus Aid, Relief and Economic Security (CARES) Act, the Consolidated Appropriations Act of 2021, and the American Rescue Plan. These measures included direct payments to individuals, expanded unemployment benefits, and support for businesses to stabilize the economy and support individuals (Gopinath, 2020). While these interventions provided much-needed relief, they also affected consumer behavior, including increased saving rates and altered consumption patterns (Muellbauer, 2021). This paper will evaluate whether consumer behavior was rational (responding to these large changes) or irrational (subject to well-documented behavioral biases). It is reasonable to expect changes in consumer behavior during an unprecedented pandemic, and this paper will evaluate whether these changes are explainable by the rational model or not.

Historical crises have similarly triggered irrational economic behaviors. For example, the 2008 housing bubble was fueled by overconfidence and herd mentality, leading to speculative property purchases and a subsequent financial crash (Shiller, 2008). Similarly, the dot-com bubble of the late 1990s saw widespread investment in internet-based companies, driven by overconfidence in their potential (Ofek and Richardson, 2003). The first recorded bubble crash was the “Dutch tulip mania”, during the 1600’s, during which contract prices for tulip bulbs reached high levels and then collapsed (Garber, 1989). It has been dubbed as a “mania” due to the *manic* behavior: tulip contracts sold for more than 10 times the annual income of a craftsman, which would be considered the same price as a luxury mansion today. People were buying tulips at higher and higher prices, intending to resell them for a profit. However, such a scheme could not last because tulip prices grew faster than the income. This showcases society's susceptibility to speculative bubbles, and irrational behaviors when these bubbles are present (Garber, 1989). This research paper will evaluate whether common behavioral biases observed in economic situations operated during the COVID-19 pandemic, a worldwide health scare that created a massive financial bubble. In this paper, I analyze behavioral biases in two distinct parts. It first explores prevalent biases seen during the pandemic, examining how they influenced consumer behavior amid social isolation, economic uncertainty, and significant government interventions (Gopinath, 2020; Muellbauer, 2021). The first section will highlight how these biases manifested in spending and investment behaviors by detailing each bias’s origins and impact on decision-making, supported by examples and empirical data. The second part evaluates biases less evident during the pandemic, exploring their usual roles and why they were less influential during this period. Considering the pandemic exceptional nature, rapid stimulus measures, and shifts in consumer priorities, this paper aims to assess the relevance and impact of these biases, aiming to understand their significance in shaping economic behavior, rather than proving their presence or absence. By examining key biases- the availability heuristic, loss aversion, herd mentality, mental accounting, and overconfidence bias- this study assesses the psychological factors that influenced consumer actions during this unprecedented period, specifically in the United States. The pandemic altered nearly every aspect of individuals' daily routines. This paper aims to assess whether common behavioral biases functioned similarly during the pandemic

“bubble” as they did before and after this period.

2. Behavioral Biases Present During the Pandemic:

2.1 Overconfidence Bias

The term “overconfidence bias” originates from cognitive psychology and behavioral economics. It gained prominence through the work of psychologists Daniel Kahneman and Amos Tversky, who conducted extensive research on human judgment and decision-making under uncertainty (Kahneman and Tversky, 1974). The overconfidence bias is a cognitive phenomenon where individuals overestimate their abilities, knowledge, or the accuracy of their predictions. Such overestimation can lead to flawed decision-making and irrational behavior. Common examples include overestimating one’s ability to predict market trends or undervaluing the risks associated with investment decisions.

During the COVID-19 pandemic, overconfidence influenced various behaviors, such as underestimating the severity of the virus and overestimating personal financial stability. The pandemic intensified overconfidence in several ways. Prolonged social isolation with economic uncertainty drove individuals to seek alternative investment opportunities, often with heightened risk. This was notably prevalent in the cryptocurrency market. The dramatic rise in cryptocurrency investments can be attributed to several factors driven by this bias. During the pandemic, many investors displayed increased overconfidence, driven by the allure of rapid financial gains in a volatile market. The perceived potential for high returns led people to underestimate the risks involved, contributing to an inflated sense of control over their financial outcomes.

According to a study by the Queensland University of Technology (QUT), the market capitalization of cryptocurrencies surged from approximately \$191 billion in January 2020 to \$769 billion by December 2020 (QUT,2021) This increase was accompanied by a 700% rise in Bitcoin prices from March 2020 to March 2021. On March 1, 2020, Bitcoin was trading at \$8,562 per coin. On March 7, 2021, there was an exponential increase as it climbed to \$51,207 per coin. This sharp rise in Bitcoin value, especially during the pandemic, reflects the impact of overconfidence bias on investment decisions. In 2022, it decreased to around \$48,086 per coin, showing that this spike was specifically during the peak of COVID-19.

Similarly, Dogecoin, a similar scheme that began as an online meme in which Internet users convinced each other that it would be amusing to coordinate purchasing a cryptocurrency, experienced an extraordinary rise in value during the pandemic. Originally launched in 2013 as a parody of the cryptocurrency hype, Dogecoin saw its value skyrocket during the pandemic, particularly in early 2021, driven by social media trends and celebrity endorsements. For instance, a single tweet from Elon Musk in January 2021 caused Dogecoin's value to surge by over 800% within 24 hours (Fortune, 2021). Investors, driven by overconfidence and influenced by social media, poured money into Dogecoin without fully understanding its speculative nature, leading to substantial financial gains for some, but significant losses for others when the bubble

eventually burst.

Another example of speculative cryptocurrency investment during the pandemic was the rise of Non-Fungible Tokens (NFTs). NFTs, digital assets representing ownership of unique items like art or music, gained immense popularity in 2021. The overconfidence bias led many to believe that digital assets would continue to appreciate indefinitely. However, this market is highly speculative, with values fluctuating dramatically. For example, the price of the famous “CryptoPunk #7804” NFT surged to over \$7.5 million in March 2021, only to see similar assets lose significant value later in the year (BBC, 2021). This speculative frenzy, fueled by overconfidence, resulted in irrational investment behaviors, with individuals pouring vast sums into highly volatile markets.

Investors, driven by overconfidence, perceived cryptocurrencies as a secure investment amidst economic instability. It led to a surge in investments without a full understanding of the market's speculative nature. The overconfidence bias during the COVID-19 pandemic led to irrational behaviors in cryptocurrency investments. Typically, overconfidence bias causes individuals to overestimate their knowledge and abilities, leading to flawed decision-making. Rational investment strategies would involve people cautiously approaching investments in volatile markets like cryptocurrencies, considering the inherent risks and potential for loss. The allure of high returns in a time of economic instability led to an underestimation of the risks, reflecting an irrational response influenced by the pandemic's unique pressures.

The widespread economic uncertainty - where roughly 9.6 million US citizens lost their jobs (Bennet, 2020) and prolonged social isolation created an environment where many individuals felt a need for control and potential financial gain. This situation was compounded by the stress associated with the pandemic. According to a study by Dreyer et al.(2021), individuals under acute stress with lower cortisol responses tend to make riskier decisions. Stress can impair judgment, making risky investments appear more attractive as a means of coping with uncertainty.

In essence, the pandemic's condition made the overconfidence bias more pronounced, as individuals sought quick financial solutions amid widespread economic disruption. This behavior was irrational, as it often disregarded the significant risks associated with such investments (such as market volatility, regulatory risks, security risks, lack of consumer protections, etc.), influenced heavily by the extraordinary circumstances of the COVID-19 pandemic. (Frontiers in Psychology, 2021)

2.2 The Availability Heuristic through Narrative Economics

The availability heuristic is a mental shortcut that relies on immediate examples that come to mind when evaluating a specific topic, concept, method, or decision. This bias operates on the assumption that if something can be recalled easily, it must be important and more frequent. It is a bias in which how you perceive the chance of an event happening is based on how easily it comes to your mind. Daniel Kahneman and Amos Tversky introduced the concept in their work on human judgment and decision-making, demonstrating how this mental shortcut

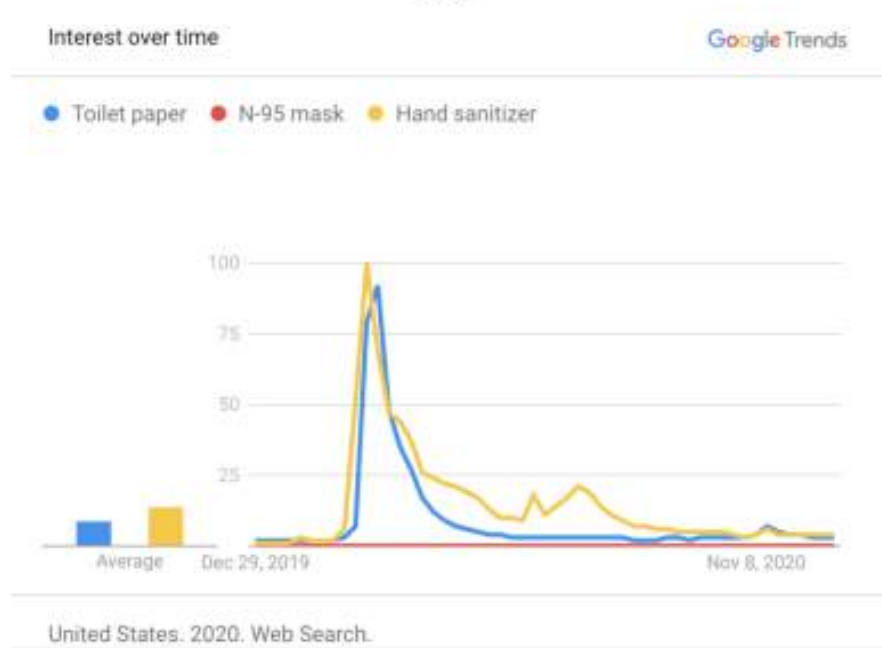
can lead to biased perceptions and irrational choices (Tversky and Kahneman, 1974). The term was coined after the ‘Linda the Bank Teller’ case, where the researchers presented participants with a description of a woman named Linda, who was said to be 31 years old, single, outspoken, and very bright, with a deep concern for social justice. Participants were then asked to rank the likelihood of various scenarios involving Linda, including whether she was a bank teller or a bank teller who was also active in the feminist movement. Despite the statistically greater likelihood that Linda was simply a bank teller, most participants chose the latter option, influenced by the vivid description that made the feminist aspect more available and thus more prominent in their minds (Tversky and Kahneman, 1983). This example illustrates how the availability heuristic can lead to biased judgments by overestimating the probability of events or characteristics that are more easily recalled or imagined, even when they are less likely to occur in reality.

During the COVID-19 pandemic, the availability heuristic significantly influenced consumer behavior. People use this heuristic when they don't have information about actual probabilities, and this was relevant during COVID, which was a very new and unknown situation. The extensive coverage of product shortages, particularly for items like toilet paper and hand sanitizer, triggered widespread panic buying, despite stable supply chains. The frequent portrayal of empty store shelves and urgent calls to secure essential goods led many individuals to prioritize these concerns disproportionately, acting on the vivid and immediate availability of these stories rather than objective data.

The availability heuristic ties in with Robert J. Shiller's concept of narrative economics. (Shiller, 2019). Shiller argues that economic events and decisions are heavily influenced by the stories that circulate in society. These narratives can drive major economic trends and behaviors by shaping perceptions and actions. During the pandemic, the narratives about scarcity and panic buying were highly influential and self-reinforcing; the stories circulating about toilet paper shortages and the associated media coverage were not just reflecting reality, but also actively shaping it.

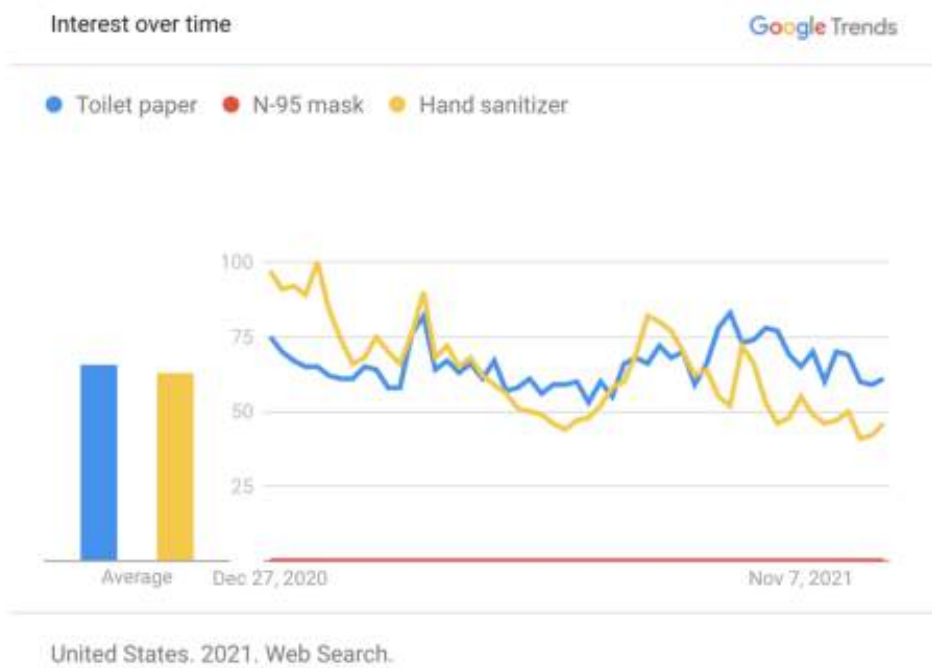
Quantifying media coverage can be challenging due to the sheer volume and variety of sources involved. However, general trends indicate that the media devoted significant coverage to issues of scarcity during the COVID-19 pandemic. According to Google Trends data, terms such as “N-95 Masks”, “Toilet Paper” and “Hand Sanitizer” saw peak search interest and viewership, with each term reaching 100% interest in various parts of 2021. The spike in searches from 2020 to 2021 reflects the intense public focus on these topics, reinforcing the narrative of scarcity (as seen in Fig. 1 and Fig. 2). The media coverage made these items appear more essential and scarce, even when supply chains remained largely intact.

Figure 1: Google Trends search interest in the terms "Toilet Paper", "N-95 Mask" and "Hand Sanitizer" during 2020.



This figure shows consistent disinterest in all three terms, except the spike in March 2020 where both Hand Sanitizer and Toilet Paper reached an almost 100% interest rate, which was the rise of the Pandemic.

Figure 1: Google Trends search interest in the terms "Toilet Paper", "N-95 Mask" and "Hand Sanitizer" during 2021.



This figure shows consistent spikes in interest for both Hand Sanitizer and Toilet Paper throughout 2021, with interest staying between 50 to 100% throughout the year.

The toilet paper shortage during the COVID-19 pandemic serves as a clear-cut example of how the availability heuristic can lead to irrational consumer behavior. Despite a robust supply chain and fully operational manufacturing facilities, the media coverage and panic-buying behaviors led to widespread shortages. In 2021, the interest in Hand Sanitizer and Toilet Paper was extremely high (Fig. 2), an exponential rise compared to the 2020 interest rates. Experts like Dr. Ronald Gonzalez pointed out that the real issue was not supply chain disruptions but rather the sudden surge in consumer demand driven by media narratives (Gonzalez, 2020). The supply chain for toilet paper was stable; the issue was people thinking that there was a shortage, and acting on that belief without statistical evidence. More than 99% of tissue products were being manufactured domestically and production facilities were operating normally (Gonzalez, 2020). However, the sudden surge in consumer demand- driven by heightened fears and media narratives- quickly depleted inventory levels. Gonzales noted that the increased demand for consumer toilet paper, due to more people staying at home, contrasted with the decreased need for commercial toilet paper, contributing to the temporary shortages. This situation highlights how the combination of media narratives and psychological biases can significantly impact consumer behavior and exacerbate perceived scarcities. (Gonzalez, 2020).

Similarly, hand sanitizer became a focal point during the pandemic, with media reports emphasizing its importance in preventing virus transmission. While this good was essential during this worldwide health issue, the narrative led to widespread stockpiling and price increases. The narrative surrounding its necessity led to widespread stockpiling and price increases. For example, the price of *Wet Ones Big One's* hand sanitizer rose significantly, reaching over \$4 per bottle, compared to pre-pandemic prices of less than \$2 (Statistia, 2020). The availability heuristic made these products appear more essential and scarce, prompting consumers to purchase them in large quantities, often beyond their actual needs. The availability heuristic made these products appear more essential and scarce, prompting irrational purchasing patterns.

Another example of this bias in action can be seen in the market for personal protective equipment (PPE) during the pandemic. N95 masks, in particular, became a highly sought-after item, with media reports frequently highlighting their scarcity and the dangers of being without them. A study by the National Bureau of Economic Research found that mask prices increased by over 500% in some regions during the early months of the pandemic, driven largely by consumer panic rather than actual supply constraints (NBER, 2021). The rationality of this is called into question because, during the early stages of the pandemic, there was a genuine scarcity of N95 masks due to increased demand, which led to hoarding behaviors. However, as production ramped up, an oversupply of masks began to occur, particularly among new manufacturers who struggled to connect with buyers. By 2021, millions of masks went unsold despite facilities still needing PPE. This suggests that panic buying initially exacerbated the issue (Sathya, 2021).

While it is hard to make a definitive causal claim that media coverage alone caused these behaviors, the timing of these consumer actions coincided with the peak of media narratives

about shortages. It would be considered rational behaviour to stockpile if shortages were legitimately happening; however, data from NCSolutions shows that while there was a shortage of toilet paper in grocery stores on April 19th, 2020, it was because of the sudden surge in panic buying, and not because of actual disruptions in supply. This is a similar case for both N95 masks and hand sanitizer. The stockpiling behavior observed during the pandemic, influenced by the availability heuristic and amplified by media narratives, deviated from rational consumer behavior. Rather than making purchasing decisions based on actual risk assessments and needs, consumers acted on vivid and repeated stories of shortages. This led to over-purchasing and hoarding, which paradoxically contributed to temporary shortages and inflated prices. Studies and surveys show that many households purchased significantly more toilet paper than usual, far exceeding their actual needs. For example, some reports indicated that people were buying dozens of rolls at once, which was not necessary for typical household consumption. Despite media reports, the supply chain was strong throughout the pandemic. The actual risk of prolonged shortage was minimal compared to the urgency created by media narratives. Data from Georgia-Pacific estimated that an average American household uses about 409 rolls of toilet paper a year. With people spending more time at home, the expected increase in consumption was around 40%. This increase, while significant, did not justify the extreme levels of stockpiling observed throughout the country. According to the rational economic model, one would look at data presented about the shortages rather than lean into publicized media coverage before stockpiling. The uncertainty of the pandemic psychologically influenced people to panic-buy.

In conclusion, the intersection of media coverage and psychological biases like the availability heuristic can significantly impact consumer behavior. The examples of toilet paper, hand sanitizer, and PPE shortages during the COVID-19 pandemic illustrate how powerful narratives can shape perceptions and drive economic behavior in ways that diverge from rational analysis and actual supply conditions. These cases demonstrate the importance of understanding how narratives influence decision-making, particularly in times of crisis.

2.3 Herd Mentality

Herd mentality, also known as mob mentality or herd behavior, refers to the phenomenon where individuals in a group adopt certain behaviors, attitudes, or beliefs simply because others are doing so, often without critically evaluating the situation themselves. This bias is driven by the desire to conform and avoid being the outlier, leading to decisions that are heavily influenced by the actions of others rather than personal judgment or rational thought. In economic terms, herd mentality can significantly impact consumer behavior, leading to trends and market movements that are not necessarily based on intrinsic value but on the perceived actions of others.

The COVID-19 pandemic created an environment where herd mentality could flourish, particularly due to the increased reliance on social media and digital platforms. With widespread lockdowns and restrictions on physical interactions, people turned to online communities for

information, social interaction, and entertainment. Social media platforms, in particular, became a powerful tool for spreading trends and influencing consumer behavior. The fear, uncertainty, and isolation brought on by the pandemic amplified the effects of herd mentality, as people sought reassurance and validation from others in a time of crisis.

Social media trends, driven by influencers and viral content, significantly impacted consumer behavior during the pandemic. The increased time spent on platforms like Instagram, TikTok, and Facebook led to a surge in online shopping, as consumers were exposed to and influenced by the purchasing behaviors of others. One concrete example of herd mentality during the pandemic was the surge in demand for home fitness equipment, largely driven by social media trends. During the pandemic, influencers popularized home workouts, leading to a surge in demand. Peloton, a high-cost option compared to other fitness solutions, experienced a dramatic increase in sales as users shared their workout experiences online. Peloton's revenue doubled in 2020 to 1.82 billion, and again to 4 billion in 2021, highlighting the influence of herd mentality (Etling, 2022). While buying a Peloton bike might seem like a rational response to gym closures, there are indicators that herd mentality played a significant role. Peloton bikes are expensive, costing between \$1,495 and \$2,495, with additional subscription fees for classes. Many consumers could have chosen more affordable alternatives (resistance bands or dumbbells), suggesting that the choice to invest in this product was influenced more by social media trends than purely rational decision-making.

The post-pandemic period also revealed overestimations in consumer demand. Peloton's active subscribers decreased significantly, and the company reported a 22% year-on-year revenue decline in 2023, which indicates a sharp drop in usage as normalcy returned (Business of Apps, 2024). This decline can lead to buyer's remorse, as seen with Peloton users who later regretted their expensive purchases (Harvard Business Review, 2023).

Another significant example was the sudden demand for certain foods and wellness products, heavily influenced by social media platforms. For instance, TikTok trends like the “Dalgona coffee” challenge and “banana bread” baking led to spikes in demand for coffee, sugar, and baking ingredients. Nielsen data from 2020 showed a 54% increase in baking ingredient sales during the early months of the pandemic (Nielsen, 2020). The herd mentality, fueled by viral social media trends, pushed consumers to participate in these activities, often leading to shortages of specific products.

Lastly, the surge in stock market investments by retail investors during the pandemic can also be attributed to herd mentality, with social media platforms like Reddit playing a crucial role. The “GameStop” stock phenomenon, driven by the Reddit community r/WallStreetBets, saw thousands of retail investors collectively driving up the price of GameStop shares in January 2021. It was essentially a herd of investors creating a bubble and trying to get rich from it, even though it was destined to fail. This event was a clear example of herd behavior, where individual investors followed the crowd without fully understanding the risks, leading to extreme volatility in the stock market (Eaton, Green, Roseman, & Wu, 2022).

Under normal circumstances, herd mentality might still influence consumer behavior, but

its effects are typically less pronounced and more localized. In everyday situations, people are more likely to have access to diverse sources of information and social interaction, which can help mitigate the impact of herd behavior. Moreover, the absence of a global crisis like the COVID-19 pandemic means that the emotional drivers of herd mentality, such as fear, anxiety, and uncertainty, are less intense.

During the pandemic, however, the combination of social isolation, increased online activity, and heightened emotional states created the perfect conditions for herd mentality to thrive. The rapid spread of information (and misinformation) on social media, coupled with the lack of alternative sources of social validation, meant that people were more likely to follow the crowd without critically evaluating the situation. This leads to consumer behaviors in various target demographics being heavily influenced by online trends, often resulting in irrational decision-making, such as panic buying or purchasing products based solely on their popularity rather than their actual utility or value.

In conclusion, while herd mentality is a prominent bias in consumer behavior, the pandemic was able to magnify its effects, leading to more extreme and widespread instances of this bias. The unique conditions of the pandemic- social isolation, increased reliance on digital platforms, and heightened emotional states created an environment where herd mentality could flourish, driving significant shifts in consumer behavior that might not have occurred in normal times.

2.4 Loss Aversion

Loss aversion is a key concept in behavioral economics, first introduced by Daniel Kahneman and Amos Tversky as part of their Prospect Theory. It refers to the tendency for people to prefer avoiding losses rather than acquiring gains asymmetrically. The pain of losing a certain amount of money is often more intense than the pleasure of gaining the same amount. This bias significantly impacts decision-making, which leads individuals to make choices that avoid potential losses. In economic terms, loss aversion can lead people to make overly conservative choices to avoid potential losses, even when the potential benefits outweigh the risks (Kahneman & Tversky, 1979).

During the COVID-19 pandemic, there was economic instability due to a lack of movement. There was a reasonable fear of financial loss which influenced saving and spending habits during the pandemic. Despite real wages rising, the persistent and visible inflation in 2022 and 2023 led many to feel a loss of purchasing power. This perception was particularly acute given the overall economic uncertainty and disruptions. As the economy began to recover and inflation surged, consumers experienced heightened anxiety about financial stability, reflecting loss aversion.

In the U.S., inflation rates from 2021 to 2023 were the highest seen in 40 years, significantly impacting the cost of living (Federal Reserve Board, 2022). Despite nominal wage growth, inflation eroded real wage gains, leading consumers to feel that their purchasing power was decreasing. The perceived loss from inflation overshadowed the actual gains from rising

wages, leading to overly cautious financial behavior. For example, people cut back on spending and increased savings, a behavior driven by the desire to avoid the perceived losses associated with inflation (Schanzenback & Pitts, 2022). The perceived loss from inflation overshadowed the actual gain from risking wages, leading to overly cautious financial behavior.

The prominence of loss aversion during this period can be attributed to the widespread job insecurity and economic instability brought about by the pandemic, a real risk. Facing potential job losses and financial distress, individuals became more conservative with their finances, saving more and spending less as a buffer against the risk of unemployment and economic downturns. With economic uncertainty and job losses affecting many, consumers were understandably cautious. This cautious behavior, while influenced by loss aversion, also reflected a rational response to genuine economic threats (Park et al., 2021)

In summary, while loss aversion influenced economic behavior during the pandemic, it was intertwined with rational responses to real economic threats. The fear of financial instability and cautious spending were logical and prudent choices rather than purely driven by psychological biases. This illustrates the interplay between behavioral biases (loss aversion) and rational financial decision-making during times of crisis. Hence, while the pandemic did see a surge of loss aversion behaviors, this can be considered rational looking at the context of the situation.

3. Behavioral Biases Not Present During the Pandemic:

3.1 Mental Accounting

Mental accounting is a cognitive bias where individuals categorize and treat money differently depending on its source, intended use, or other subjective criteria. This can lead to irrational financial decisions as people may not always allocate their resources in ways that maximize their overall utility. Typically, in a normal economic setting, mental accounting manifests as people treating money from different sources (e.g., bonuses, gifts) differently from their regular income. For instance, someone might spend a tax refund on a vacation while using their regular salary to cover daily expenses, even though both funds are ultimately interchangeable. The term, coined by economist Richard Thaler, is a way to describe how people irrationally distribute their money (Thaler, 1999).

During the pandemic, mental accounting appeared less prevalent or manifested differently, and that can be attributed to the extraordinary economic and psychological pressures. The unprecedented nature of the pandemic and its impact on financial stability led many individuals to prioritize immediate needs over discretionary spending, regardless of the source of the funds. This is clear through the distribution of stimulus checks in the U.S.

The USA's distribution of stimulus checks during the pandemic provides a clear example of mental accounting in action. As of early May 2021, the federal government had enacted legislation to provide about \$850 billion in Economic Impact Payments to taxpayers (U.S. Department of Treasury, 2021). These payments helped mitigate the loss of employment-based

income for many households and temporarily eased the economic damage from the pandemic. However, the way Americans spent that money differed across each round of payments, influenced by mental accounting.

There were three main rounds of stimulus checks, and it was spread out differently. One might expect people to treat stimulus payments as “bonus money” or separate from regular income, but the overwhelming economic stress caused many to use these funds primarily for essential expenses rather than treating them as separate or additional funds. (Chetty et al., 2020)

Figure 3: Overview of Stimulus Payments

Round Number	First Round (CARES Act)	Second Round (Consolidated Appropriations Act)	Third Round (American Rescue Plan)
Amount	\$1,200 per adult, \$500 per child	\$600 per adult, \$600 per child	\$1,400 per adult, \$1,400 per child
Income Threshold	Phased out for single taxpayers earning above \$99,00, heads of household above \$136,500, and married couples above \$198,00	Phased out for single taxpayers earning above \$87,00, heads of household above \$124,500, and married couples above \$174,00	Phased out for single taxpayers earning above \$80,000, heads of household above \$120,00, and married couples above \$160,000

Cost \$292 billion \$164 billion \$411 billion Information gathered from U.S Department of Treasury (2021 Report); Economic Impact Payments

Mental accounting led many recipients to treat the stimulus checks as separate from their regular income, often resulting in different spending behaviors. A rational consumer would ideally use this unexpected income to address their most pressing financial needs, such as paying down debt, covering rent, or building emergency savings. However, the categorization of stimulus checks as “bonus money” influenced people to spend on non-essential items rather than prioritizing urgent financial needs (Parket et al., 2020)

The first round of stimulus checks, sent through the CARES Act, was disbursed in April 2020 when unemployment was at its highest (14.7%). Nearly three-quarters of households used or planned to use the \$1,200 payments primarily for expenses like food and rent. This spending was seen as necessary given the immediate financial strain. The impact of this was that it created an immediate boost to consumer spending, helping to stabilize the economy temporarily. According to the Bureau of Economic Analysis (BEA), approximately 74% was spent on Spending, 14% on Saving, and 11% on paying debt (BEA, 2020). Many recipients used their stimulus checks for essential expenses such as food, rent, and utilities. This reflected a prioritization of survival needs over discretionary spending, demonstrating a shift from typical mental accounting practices where such funds might be categorized differently. At the time, the pandemic was a disease with a lot of unknown variables, and therefore the uncertainty led to

people spending money to protect themselves in the present. This can be considered rational behavior, since in uncertain times of crises, it is considered prudent to focus on the present rather than saving and paying debt for the future.

The second and third round of spending was disbursed in January and March 2021, with some economic recovery underway. There was a significant shift towards saving and paying off debt, with only about one-fifth of households reporting using the payments primarily for expenses (Ciobion et al., 2020). There was an increase in personal saving rates (e.g., saving rates rose to 20.0% in January 2021) and debt reduction. This behavior was less stimulative in the short term but potentially beneficial for long-term financial stability. In the second stimulus check, 22% was on spending, 26% on saving, and 51% on paying debt. The shift towards saving and debt repayment indicates that as the economic situation began to stabilize, people started viewing the stimulus funds as an opportunity to improve their financial security. For example, many households paid down credit card debt or contributed to saving accounts, demonstrating a change in mental accounting from immediate spending to future financial planning (BEA, 2021). Similarly, in the third and final stimulus check, 19% was on spending, 32% on saving, and 49% on paying the debt. The trend continued where, instead of treating the funds as additional resources, it was being incorporated for broader financial planning. The majority was used to save or pay off debt. Households earning more than \$150,00 annually were more likely to save or pay off debt compared to spending on essential expenses, showing that higher-income individuals treated the checks as an addition to their financial security rather than immediate spending needs (Park et al., 2021).

From a behavioral economics point of view, it was to be expected that people would view stimulus payments as separate from regular income and potentially spend them on non-essential items. However, the actual behavior shows that the economic strain led many to use the funds for immediate needs, reflecting a blending of mental accounts. One would think that individuals might categorize income from work and government aid differently, but economic pressures blurred these lines.

In a pre-pandemic economic situation, mental accounting often leads to distinct spending behaviors based on the perceived source of funds. For example, people might view tax refunds or gifts as extra money to be spent on luxuries or non-essential items, while their regular income is reserved for necessary expenses. During the pandemic, however, the immediate financial needs caused many to blend different sources of funds, using stimulus payments for urgent necessities like food, rent, and utilities. This deviation from typical mental accounting practices reflects how extreme circumstances can reshape financial decision-making processes. The pandemic led to a more unified approach to spending, where the perceived purpose of funds was overshadowed by immediate financial needs.

4. Discussion

The COVID-19 pandemic provided a unique environment for examining behavioral biases in economic decision-making. The analysis reveals that certain biases, such as herd

mentality and loss aversion, were prominent during the pandemic, while others, such as mental accounting, were less apparent. The variability in the manifestation of behavioral biases during the COVID-19 pandemic can be attributed to several factors, including the severity of the crisis, the nature of decision-making under stress, and the quality of information available. Understanding these factors can provide insight into why some biases are more prominent than others.

One potential explanation for the presence or absence of certain biases during the pandemic is the concept of bounded rationality. Simon (1975) introduced bounded rationality to describe the limitations on individuals' decision-making capabilities due to cognitive constraints. During a severe crisis, such as the pandemic, the overwhelming seriousness of the situation may lead individuals to make more rational decisions, focusing on immediate survival and practical needs rather than engaging in biases like mental accounting. For instance, the need to secure essential goods and maintain financial stability could have driven more uniform decision-making processes, diminishing the typical effects of mental accounting, where individuals categorize and allocate funds for different purposes (Simon, 1957; Thaler, 1985).

Additionally, the severity and nature of the pandemic created an environment where certain biases were either suppressed or amplified. The urgency of the crisis likely led people to rely on simpler, more instinctual decision-making processes. For example, while herd mentality and loss aversion are more instinctual reactions to perceived threats and uncertainty, mental accounting requires a more deliberate and segmented approach to financial decision-making. The pandemic heightened the prominence of certain biases, such as herd mentality, loss aversion, overconfidence bias, and the availability heuristic. One major factor contributing to this amplification is the increased uncertainty and perceived threat posed by the pandemic. In times of crisis, when reliable information is scarce and fear is prevalent, people tend to rely more heavily on mental shortcuts or heuristics to make decisions. The reliance can intensify biases like herd mentality, where individuals look to others for cues on how to behave, leading to collective actions like panic buying and stockpiling (Chou, 2021; Smith, 2023). Loss aversion was amplified due to uncertainty about the future, making individuals more risk-averse. Similarly, the overconfidence bias was also evident and is often driven by an individual desire for control and optimism, which can become more pronounced in uncertain times as a way to cope with fear and anxiety (De Bondt & Thaler, 1995).

On the other hand, biases like mental accounting were less prominent, likely due to the overwhelming nature of the crisis. The pandemic forced individuals to focus on immediate survival needs, leading to a more unified and less segmented approach to financial decision-making. The urgency of addressing basic needs overshadowed the typical mental categorization of funds, as people blended different financial resources to navigate the crisis effectively (Thaler, 1985; Parker et al., 2021).

Understanding which biases emerge during a crisis can inform strategies for improved decision-making in future scenarios. Biases like herd mentality and loss aversion often become more pronounced during times of uncertainty and rapid information dissemination, while biases

related to resource categorization, such as mental accounting, may be less influential when immediate survival and practical needs take precedence. To better navigate economic challenges, future crises may benefit from strategies that address the emotional and psychological drivers of these biases, such as providing clear, reliable information and fostering rational decision-making frameworks. By recognizing the interplay between cognitive biases and situational factors, policymakers and individuals can work to mitigate the effects of behavioral biases and make more informed choices.

5. Conclusion

This study explored how behavioral biases influenced consumer behavior during the COVID-19 pandemic, focusing on herd mentality, loss aversion, mental accounting, and their interactions. We also evaluated the presence of biases in the context of pandemic-specific economic pressures.

Herd mentality emerged strongly due to heightened social media usage and digital trends. Consumers were heavily influenced by online behaviors, leading to trends such as panic buying and the surge in the popularity of certain products. Loss aversion was evident as the economic instability and inflation led people to be more conservative with their spending and increased savings. The psychological impact of perceived financial loss overshadowed nominal wage gains, reflecting cautious financial behavior. Overconfidence bias was present in cryptocurrency investments, where investors underestimated risks despite volatile market conditions, driven by optimistic expectations. The availability heuristic influence consumer decisions, as the prominence of certain products and information during the pandemic led to skewed perceptions of necessity and value.

Mental accounting, however, was less noticeable during the pandemic. The urgent financial needs caused by the crisis led people to blend different sources of funds, using stimulus payments for essential expenses rather than treating them as separate "bonus" money. In summary, while some biases like herd mentality and loss aversion were pronounced due to the unique economic and psychological conditions of the pandemic, others like mental accounting were less relevant. Understanding these dynamics offers insights into how biases may shift in future crises and highlights the need for strategies to manage their impact effectively.

Works Cited

Journal Articles

- Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data. National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w27431/w27431.pdf
- Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). How did U.S. consumers use their stimulus payments? National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w27693/w27693.pdf
- Goolsbee, A., & Syverson, C. (2021). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, 193, 104311. <https://www.sciencedirect.com/science/article/pii/S0047272720301754?via%3Dihub>
- Gopinath, G. (2020). The great lockdown: Dissecting the economic effects. IMF Blog.
- Muellbauer, J. (2021). Pandemic economic uncertainty and saving: Theory and evidence. *Review of Economic Dynamics*, 40, 68-101. <https://www.sciencedirect.com/science/article/abs/pii/S1094202521000326?via%3Dihub>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185 (4157), 1124-1131. <https://www.science.org/doi/10.1126/science.185.4157.1124>
- Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 47 (2), 263-291. <https://www.jstor.org/stable/1914185>

Books

- Garber, P. M. (1989). Tulipmania. *Journal of Political Economy*, 97 (3), 535-560.
- Shiller, R. J. (2008). *The subprime solution: How today's global financial crisis happened, and what to do about it*. Princeton University Press.
- Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.
- Simon, H. A. (1957). *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*. Free Press.

Other Sources

- Federal Reserve Board. (2022). Consumer price index, inflation, and wages. Federal Reserve Economic Data. Retrieved from <https://www.federalreserve.gov/>(<https://www.federalreserve.gov/>)
- National Bureau of Economic Research (NBER). (2021). The impact of media coverage on consumer behavior during COVID-19. NBER Working Paper No. 28912. Retrieved from <https://www2.nber.org/papers/w28912>](<https://www2.nber.org/papers/w28912>)
- Parker, K., Minkin, R., & Bennett, J. (2021). Economic fallout from COVID-19 continues to hit lower-income Americans the hardest. Pew Research Center. Retrieved from

<https://www.pewresearch.org/social-trends/2020/09/24/economic-fallout-from-covid-19-continues-to-hit-lower-income-americans-the-hardest/>

Schanzenbach, D. W., & Pitts, A. (2022). How inflation is eroding real wage gains. *Science Direct*. <https://doi.org/10.1016/j.jpolmod.2022.03.001>

Statista. (2020). Price of hand sanitizer during the COVID-19 pandemic. Retrieved from <https://www.statista.com/statistics/1050783/average-price-of-leading-hand-sanitizer-brands-us/>

IMF. (n.d.). Policy responses to COVID-19. International Monetary Fund. Retrieved from <https://www.imf.org/external/error.htm>

Additional Articles and Sources

Chou, S. (2021). Social media trends and consumer behavior during the COVID-19 pandemic. *Frontiers in Psychology*, 12, 1234. <https://doi.org/10.1093/jcr/ucaa013>

Frontiers in Psychology. (2021). The psychological impact of overconfidence bias on financial decision-making during the COVID-10 pandemic. *Frontiers in Psychology*, 12, 1234.

Google Trends. (2020). Search interest in "toilet paper shortage" and "hand sanitizer." Retrieved from

<https://trends.google.com/trends/explore?date=2020-01-01%202020-12-31&geo=US&q=%2Fm%2F09gtd,%2Fg%2F11jh7gqjq,%2Fm%2F0b23px&hl=en>

Decoding Moche Civilisation: The Significance of Stirrup Spout Bottles

By Avani Radheshyam

Introduction

A stirrup-spout bottle or vessel is a ceramic vessel that was created by the Moche people in Peru between 100 and 800 CE. Its spout shape, which mimics a riding saddle strap, gave rise to its name. It has a bulbous body which holds the liquid and from the top of the body, two tubes from either side rise and join at the top, thus forming a single spout. The two tubes joining together with the spout at the top also acts as the handle for the bottle.

All of the bottles exhibit a high degree of technical and artistic proficiency. This bottle is made of terra-cotta (Fig. 1). The majority of the bottles are of the regular shape with paintings on them, while the rest of the bottles are elaborated into three-dimensional sculptures. They depict various things, featuring scenes from mythological scenarios to everyday life. The chambers of the stirrup spout bottle are molded to replicate human heads, seated figures and plants or animals. These forms are unique to Andean cultures and very complex (Donnan, 1965). Several pieces of evidence suggests that the stirrup spout bottle was an extremely large part of the Moche civilization and it certainly had several uses in their day to day life which we might still be unaware of.

When studying the Moche civilization, in order to truly understand their culture, their daily lives, their government and much more, it is important to investigate both, the literary and non-literary works. If scholars only study the literary aspects of history such as literature documents, poetry, plays and other types of written work, then it would give them a certain degree of understanding as to a society's language, education and interests. However, studying only literary aspects would also leave several gaps in such a scholar's understanding of a civilization and their history.

Historians have had very little written material to examine because the Moche people did not speak a written language. More often than not, in history, the greater part of the population did not know how to read and write which means they could not have produced enough literature for us to truly understand them. In art historian Jules Prown's words:

[Objects] offer the possibility of a way to understand the mind of the great majority of nonliterate people, past and present, who remain otherwise inaccessible except through impersonal records and the distorting view of a contemporary literary elite (Prown, 1982).

While it is true that literary work must be studied, it is extremely important to look at other aspects as well. Many historians have brought up questions against simply just studying literature. Historian and Pulitzer prize winner Laurel Thatcher Ulrich, for example, brings up important questions such as "How can we approach aspects of the past that written words do not record?" (Tangible Things." Harvard University, 2023).

This is where material culture comes into the picture. Material culture is essentially a branch of cultural history however, it is too diverse to constitute a single field. Material culture is the study of artifacts through which we can understand ideas, beliefs, habits and culture of a

particular community at a given time (as explained by Yale University professor, art historian Jules Prown) (Prown, 1982). Focusing on a singular object can generate excitement, prompt historical curiosity and produce understanding that wouldn't have risen without studying the object itself as proven by the recent Tangible Things Project at Harvard University ("Tangible Things." Harvard University, 2023).

Material culture should be used while examining the Moche people in order to fully understand them and the way they lived. The Moche people have several objects that are worth examining; however, the most iconic vessel they made is the stirrup spout bottle. For example, if we looked at the stirrup spout bottle with a cat from the collection of the Metropolitan Museum of Art, there is a lot we can understand about the Moche by examining it (Fig. 3). We could examine the material it is made out of, whether a mold was used, what the shape signifies, the cat's stance, specific features and much more which will definitely provide insight into the Moche. The stirrup spout bottle with cat will be looked at in depth later on in the paper. This paper aims to highlight the importance of studying materials in understanding history by analyzing a specific stirrup spout bottles and exploring the various meanings behind it in order to get a thorough understanding of the Moche civilization (their daily life, habits, culture, etc)



Fig. 1



Fig. 2



Fig. 3

Moche Civilization

General background

The Moche are an Andean civilization that were present from 100-800 CE in what is now known as Peru. The name "Moche " was derived from the great site of Moche and the Moche river valley, where they are thought to have resided ("Moche." Encyclopædia Britannica, Encyclopædia Britannica, 2024).

The Moche people were spread over an area from the Huarmey Valley in the South to the Piura Valley in the north. The capital is known as Moche and it lies at the foot of Cerro Blanco Mountain. As historian of ancient Andean art, Margaret Jackson asserted, Moche was never a unified empire and it is wrong to assume so. They were a collection of interdependent autonomous polities located throughout the North Coast valleys. Although they were politically independent groups, they shared common ideologies, religious practices, iconography, daily habits and governing strategies (Jackson, 2018).

Numerous historians have pondered the functioning of Moche social classes. Historians such as archaeologist Claude Chapdelaine have theorized using archeological research that the Moche were split into 4 classes: the elite class, the urban class, the middle class and the lower class. The elite and urban class resided in the *urban nucleus* (The urban nucleus is located at the geographical center of the neighborhood similar to how a nucleus is located at the center of the cell., 2025) and mostly consisted of state council members, noble families and leaders of city quarters while the middle and lower class resided outside of the urban nucleus. The middle class consisted of craftsmen and bureaucrats who lived around the urban sector. The lower class was majorly made of food producers who did not live in the urban sector at all (Claude, 2001).

The Moche are best known for their artworks that display excellent technical skills. Moche architects and artists designed monumental pyramids, temple complexes, exquisite ceramic vessels and beautiful jewelry which prove their highly skilled artistry and complex figural imagery (Jackson, 2018). Their vessels, specifically the ceramic ones, are their most unique form of artistry. They depict a variety of things such as supernatural beings, animals and plants. These vessels are painted or shaped so that they depict scenes of the Moche people's daily lives. Scenes such as fishing, religious ceremonies, hunting, commonly seen animals in their valley and relationships between people are portrayed on these vessels. (Fig. 4)

Material culture of the Moche

Max Uhle is renowned all over the world and popularly known for being the very first archaeologist to excavate the Moche site at Huaca del Sol, also known as the Moche valley in the early decades of the 20th century. His excavation brought Moche ceramics and artistry into the picture, specifically their popular stirrup spout bottle which expressed high levels of artistry and provided many insights into their lives. He discovered certain tools and textiles they used as well.

However, it was Rafael Larco Hoyle who further excavated and wrote extensively on his interpretations of what Moche was. He purchased and categorized the ceramic work of the Moche people. He believed they were ruled by religious and military authorities (Quilter, 2021). The majority of the excavation sites that found Moche remains are found in the site of Huacas de Moche on the northern coast of Peru. On this site, there are two large structures present that archaeologists believe to be two of their temples: Huaca del Sol, temple of the sun and Huaca de la Luna, temple of the moon (Huacas Del Sol y de La Luna at 'Moche Site': The Excavation Area of the, 2024).

There has been a lot of archaeological evidence found in these sites such as the monumental architecture mentioned in the previous paragraph, ceramics, burial sites, textiles, murals and other forms of artwork. The reason why we have so much knowledge about them is because of these artifacts found when excavating these sites. These artifacts depict various aspects of their lives.

Despite the fact that we have a fair amount of archaeological data on the Moche, this civilization did not leave behind extensive written records since they didn't have a written language. Their material culture is what gives almost all of our knowledge about them. Since our knowledge of them is so dependent on our understanding of their objects, there are certain limitations that come with simply just studying the objects.

One possible drawback of using these artifacts to learn about Moche civilization is that interpretations of them may be biased because there aren't enough written sources to support them. There is no single correct way to interpret the object's usage or meaning, making it easily misinterpreted and possibly leading to false conclusions about the Moche civilization.

One such limitation, specific to Moche archaeology was pointed out by Curator of the Arts of the Americas, Erell Hubert, who said that when categorizing objects, there's a divide between the objects that are considered "ritual/symbolic" and the objects considered "domestic." Doing so might overlook the true impact that these "domestic" objects had on political and social relations of the people. She argues that the deeper meaning and impact of these objects must be further studied rather than simply focusing on the practical use of it (Erell, 2016).

Through the excavation of Moche burial sites, a lot has been learnt about them due to the presence of many such objects. In fact, the majority of the artifacts related to the Moche were found in their burial sites. The tombs with the most objects and ornaments are believed to have been for the elites. These burial chambers contained several objects such as ornaments and specifically, stirrup spout bottles with various elaborations on them. The highest-status burials found to date at the site of Moche were found on the Huaca (Claude, 2001).

The different types of burials for different people proved that the Moche indeed had social classes. The burial chambers were larger for the elite class and smaller for the middle and lower class. The objects found in these burial sites that historians study today were surely made by the Moche potters, belonging to the middle class of the Moche social hierarchy.

Moche Potters

Moche potters were the artisans of the Moche civilization. They are the ones who created all of the pottery found that contributes largely towards our knowledge of them. They were extremely skilled craftspeople who molded the vessel to resemble an animal, human or plant and made realistic portraits on these vessels as well. Warriors, combat scenes and images of captives are themes that run in Moche art. Some ceramics showcase human sacrifice as well.

According to Chapdelaine, the potters belonged to the middle class who lived in the urban sector. Although they didn't reside in the elite's residential area of the city (urban nucleus), it was nevertheless rather close to it (Claude, 2001). The potters occupied a fairly low position in

the social hierarchy when compared to the noblemen, however, they played an extremely large role in providing various forms of vessels for both daily and religious use.

The Moche ceramics are some of the most advanced of their time. These ceramics are made by mixing clay, water and earthen elements. There are several possible working methods that the Moche potters could have used during their production. While these working methods will be discussed in further detail later on in the paper, here is a brief overview of how Moche potters likely made their ceramics.

One of these working methods is molding. Molding was achieved by either shaping the vessel by hand or by pressing moist clay into a mold, as American archeologist known for his work on the Moche civilization, Christopher B. Donnan theorized. The dry clay is then removed from the mold after it has achieved the desired shape/form. Coiling is used in conjunction with molding for certain parts of the vessel.

Another possibility is stamping. This occurs when the Moche potters simply press depictive figures onto the already shaped object which leaves an impression. The stamp is an addition to it. This is different from molding, where the shape of the object is changed (Donnan, 1965). These methods will be further discussed in the coming sections.

These methods of working resulted in various different types of vessels being made by the Moche potters. Art historian and scholar Lisa Trever discussed several of them in her paper, “A Moche Riddle in Clay: Object Knowledge and Art Work in Ancient Peru.” Some of them are:

- Bichrome jars and bowls have two different colors with minimal decoration (Fig. 5).
- Sculptural vessels include vases with elaborate shapes and imagery such as a bottle in the form of a skeletal potato head (Fig. 6).
- Transmogrifying vessels depict transformations such as a contorted human face transforming into a large owl.

Among these, one of the most iconic forms of vessels created by the Moche potters is the *stirrup spout bottle* (Lisa, 2019). Technically stirrup spout bottles can be classified as sculptural vessels due to their elaborate forms, however they are their own category due to their unique shape.



Fig. 4

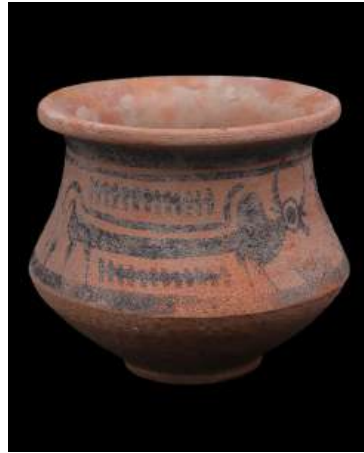


Fig. 5



Fig. 6

Stirrup spout bottle

Characteristics

The stirrup spout bottle has a variety of forms. As mentioned earlier, these vessels are made of terracotta. They usually have symmetrical chambers, which are then elaborated by stamping on the outside; however, there are also bottles with nonsymmetrical chambers. These bottles come in a variety of forms.

Stirrup spout bottles all vary amongst each other in shape, size, and decoration, however, they all have the same basic structure, with three main parts:

1. The chamber of the stirrup spout bottle is the main body where the water is stored. The outside of the chamber is where the main artwork and molding take place. The chamber is molded into either a naturalistic shape (animals and plants), or an abstract form (mythological beings). The shape of the vessel (mainly the chamber) slows the evaporation of the water ("Classroom Activity: Stirrup Bottle." Virginia Museum of Fine Arts, 2024).
2. The base of these bottles is rounded and slightly flattened. It is smooth and less decorated when compared to the rest of the bottle in order to provide support. It is blended seamlessly into the body of the bottle. The surface is usually standard, however, depending on how elaborate the figure is, it can differ.
3. The spouts of the bottle are what make it stand apart from vessels made in both Moche and non-Moche civilizations. Two spouts from either side emerge from the bottle and join at the top to create a singular opening. It efficiently pours the liquid without spilling and can also be used as a handle for the bottle. The dual spout resembles a riding stirrup, which gave rise to its name (Donnan, 1965).

Although these three structural parts have been identified, the true processes employed to create the stirrup spout bottle have never been proven, despite much debate among historians.

Methods of production

Several methods of production for the stirrup spout vessel have been pondered upon by historians. One American archeologist known for his work on the Moche civilization, Christopher B. Donnan, revolutionized our understanding of the methods of production by experiments in which he made the bottle and, based on his results, suggested what he thought was done by the Moche. Based on the conclusions from his experiments, Donnan believes that the making of the stirrup spout bottle involved four basic techniques: coiling, molding, modeling and/or stamping. According to his observations, he concludes that all these vessels require at least two of these techniques to achieve their form. He mentions that all traces of fabrication were removed during the finishing process so this theory of his is not infallible (Donnan, 1965).

He starts by saying that molding is the most basic technique. This is when moist clay is pressed into the mold and when the clay dries, it is then removed, thus achieving its desired shape/form. These molds that the moist clay is pressed into, were likely to have been made of fired clay that was formed over an actual object such as an ear of corn or squash. Donnan himself created a successful bottle by pressing clay around the mold and removing the two halves when they dried, which is how he concludes that it was likely how the Moche potters made their stirrup spout as well. (Fig. 7)

The chambers of the stirrup spout vessel are most likely to have been made using two piece molds. This means that each of the two molds were made separately and were later joined together to form the chamber of the stirrup spout. Donnan's hypothesis of having a two piece mold can easily be true because the stirrup spout bottles are extremely complex forms and are heavily detailed which cannot be easily achieved by a single piece mold. The two piece mold was likely joined using the coiling method which is why it is hard to see the separation between the two on the finished bottle. (Fig. 7)

Donnan hypothesizes that coiling was used in conjunction with molding to produce some parts of the vessel. Coiling formed the chambers of the bottle. It's usually used to form the base and join the spouts to the body of the vessel. These coils are extremely hard to detect on the bottle as they were smoothed out during production. However, coiling is one of the only logical ways the Moche could have joined one part of the bottle to the other without other obvious markings.

However, smoothing was not the only method of covering markings made by coiling. Stamping was also occasionally used.. Unlike molding, stamping is a form of surface ornamentation rather than changing the form of the object itself (Fig. 8). Donnan believes this method was used in stirrup spout bottles; most commonly when stamps, made of fired clay (each piece depicting a particular form such as an animal or human) were pressed against the inside of the vessel leaving an impression on the outside. Since this method requires the potter to reach the inside of a vessel with extremely small openings, it had to have been done before the vessel was

completed. This stamped evidence further proves Donnan's theorization of the two piece mold, because it would have been far easier for stamping to occur when the vessel was in two separate pieces, prior to sealing them together. Typically, stamping was applied to stirrup spouts with a regular shape. This essentially suggests that it was solely intended to be ornamentation added to a regular shaped stirrup spout bottle.

Modeling is a secondary technique, however, to date no Moche ceramic has been found that was made exclusively using the modeling technique. It was simply employed as a means of elaborating the object. In some cases, modeling shapes clay which is added to the bottle after its initial production; in other cases, modeling is used for re-shaping the clay that was already a part of the object. Donnan believes that modeling is how the handles and elaborations of the figures such as arms, legs and ears were made. The Moche stirrup spout vessel figures have several complex features that could not have been achieved by Moche molds Modeling is the only logical way in which these effects could have been achieved. (Fig. 9)

Donnan theorizes that these are the methods that were used by the Moche during the production process. However, his work is speculative (as he mentions in the beginning of his paper) as we cannot be completely sure whether these methods were truly used by the Moche potters or not (Donnan, 1965). While it is challenging to not know for sure how the stirrup spout vessels were produced, Donnan's theory presents a convincing argument for a set of possible techniques. Such speculation—informed by rigorous experimentation—is necessary in the field of ancient and material history, because it is extremely hard to conclude precisely how an object was made without additional evidence.

Everyday Usage

Based on the various archaeological findings and research done by historians, the stirrup spout bottle was originally made by the Moche potters for domestic use. It was mainly used in the domestic setting for storing and transporting liquid easily.

The two spouts coming from either side of the bottle connecting to form one singular spout is based on everyday use and environmental factors. In the Northern coast of Peru, where the Moche people resided, it is hot almost all year long which means water conservation must have been extremely important to them. The shape and design of the stirrup spout bottle is made specifically in order to minimize the evaporation of the liquid inside. The two spouts also act as a handle for the bottle, thus making it easy to carry around. The spouts are extremely narrow and prevent contamination, which signals it could have been used to store medicinal liquids as well; however, it must have been mostly used for domestic purposes in households.

As anthropological archaeologist Jeffrey Quilter talks about in his paper published in the *Journal of Andean Archeology*, the form of the stirrup spout bottle with a narrow spout lip makes it awkward to pour liquids. It is possible that the Moche people used stoppers in order to prevent the liquid from spilling out; however, if the stopper was made from organic material, it might have disintegrated due to poor organic preservation in the Huacas De Moche. If they did use stoppers, we have no evidence of it. It is simply a speculation (Quilter, 2021).

The stirrup spout bottles were not just used for domestic or household purposes. They were also employed in religious contexts. In fact, there is more evidence proving the use of stirrup spout vessels in the religious context than in the domestic. On these religious or ceremonial bottles, various deities, religious stories and mythological tales are portrayed through molding or slip painting (Slip paint is a liquid mixture of clay and water, which is applied to the surface of the stirrup spout bottle before it is fired, 2025). The stories the bottles convey have contributed greatly to scholars' understanding of the Moche people and their practices.

It is believed that stirrup spout bottles not only depicted the religious practices but the bottles themselves were also actively used in them. In fact, large numbers of stirrup spout vessels have been found in elite burials on Peru's northern coast and each one of them display unequaled technical and artistic skill ("Stirrup-Spout Bottle: Cupisnique." The Metropolitan Museum of Art, 2024). Historians believe that the Moche people buried their dead with these bottles and other objects because they believed that the dead would take these objects with them in the afterlife. The stirrup spout bottle was an offering to the afterlife, which provided sustenance to the deceased. The number of objects buried with a person depended on their social hierarchy. The royals were buried with gold, silver and lavish goods, but almost all of them were also buried with a specific stirrup spout bottle (Jackson, 2018).

What was it about the stirrup spout bottle that made it so significant to the Moche people—significant enough to be included with gold, silver, and other lavish goods in their funerary offerings? One possible answer is its unmistakable form. The uniqueness of this structure can truly be seen when the stirrup spout vessel is compared with pottery from other civilizations. For example, a recent exhibition at the J. Paul Getty Museum in Los Angeles, titled *Picture Worlds: Greek, Maya, and Moche Pottery* juxtaposed pottery from these three civilizations (Picture Worlds: Greek, Maya, and Moche Pottery." Getty Museum, 2024). When looking at vessels made in ancient Greece, the pottery's shape is difficult to differentiate from forms used by other civilizations. Although ancient Greek pottery does have certain distinctive features that make it unique to Greece—such as the way it has been painted—many other civilizations have produced containers with a similar shape and form due to the use of widely adopted pottery techniques. The Moche people's stirrup spout bottles are an exception to this tendency. The stirrup spout bottle stands out in global pottery traditions due to its incredibly distinctive form, which was not adapted by any other civilization during its time.

Possible meanings of the form

The distinctive shape and function of the stirrup spout bottle with the two spouts rising from either side of the bottle to form a singular spout has raised several questions about the deeper meanings behind the form. One of the theories formed about the stirrup spout bottle is that the two parts of the stirrup joining into one spout reflects Andean ideas of dualism. Dualism is the idea that in any field or domain, there are bound to be two fundamental principles that interact from time to time. For example, in religion, the two basic fundamental principles are good and evil (Robinson, 2020).

In the past, many people such as French philosopher René Descartes and ancient Greek philosopher Plato have talked about dualism. Rene Descartes believed that our reality consists of mental substance, which is the mind and physical substance, which is the body. Plato's example of dualism was the body and the soul. He theorized that the soul, the true essence of a person, uses the body as a vessel. In each of these theories, both entities are fundamentally different but they still interact with one another as they are a part of the same domain (Robinson, 2020).

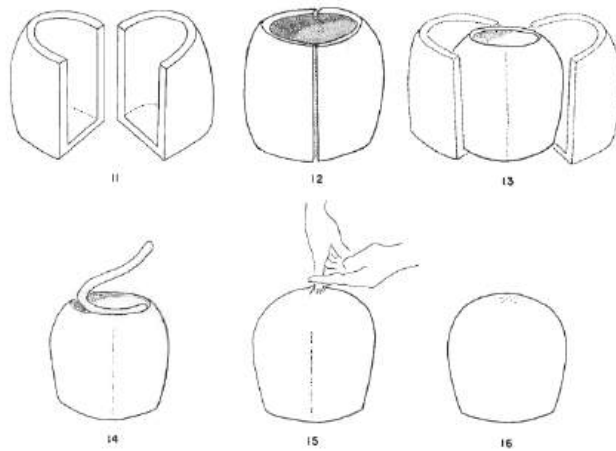
Asymmetric dualism is the idea that two entities in the same domain are imbalanced but still dependent on one another. For example, Plato's philosophy of the body-soul dualism proposes that the soul is of more value than the body. In this respect, it is an asymmetric dualism. According to research, this type of dualism was most popular amongst the ancient Andes (Quilter, 2021). However, the symmetric shape of the stirrup spout bottles may suggest that the Moche people favored balanced dualism. Balanced dualism is the opposite of asymmetric dualism. It is the idea that the two opposing entities in one domain are equal to one another, like yin and yang in Chinese philosophy (Quilter, 2021).

If we analyze a basic stirrup spout bottle, the two spouts on either side may stand for the two opposing entities in the same domain. Since they emerge from either side of the same chamber, the domain in this instance would be the chamber from which they both arise. When the two spouts come together to form a single spout at the top of the bottle, that might stand in for the inevitable interplay between opposing entities found in the concept of dualism.

Moche people had their own dualist philosophies. This dualism, known as the concept of *tinku*, has been identified by archaeologist Michele L. Koons. Tinku is an ancient Andean concept of two parts coming together to create a whole. It offers insight into how the Moche viewed the world. In Quechua, an ancient indigenous language family originating in Peru, the word *tinkuy* quite literally means binding or meeting. It can also be translated as to encounter or to come together. Tinku implies that the two opposing halves achieve equilibrium. Koons uses the opposition between a male and female pair as an illustration of how, despite their differences, they must work together to form a whole in order for society to operate. The two separate channels of the stirrup spout bottle eventually uniting into one spout is the perfect embodiment of the concept of tinku (Koons, 2022).

Another theory behind the stirrup spout bottle was brought up by historian and Pulitzer-prize winner Laurel Thatcher Ulrich. In an essay on the Tangible Things Project at Harvard University, Ulrich conducted a brief analysis of a stirrup spout bottle which depicted a mother holding her child. According to her theory, this vessel, along with many of the stirrup spout bottles, signifies how water moves through human bodies, which represents the Moche society's reliance on their highly developed irrigation system (Tangible Things." Harvard University, 2023). Inspired by Koons, Ulrich, and others, this paper seeks to identify the meanings behind the stirrup spout vessel's form. There is a clear relationship between the Moche surroundings—in the case of Ulrich's vessel, the irrigation system—and the forms/motifs on the bottles. Many bottles feature depictions of the natural world, including animals. What are some

possible meanings behind these forms, and how do they provide insight into the Moche's relationship with animals?



(Donnan, 1965)

Fig. 7



Fig. 8



Fig.

Case study



Fig. 10

This is a stirrup spout bottle representing a pampas cat that is currently being stored in the Metropolitan Museum of Art (fig. 10). The pampas cat is a small wild cat, native to South America. It is named after the low South American grasslands Pampas (Horizonte, Prefeitura de Belo. “Pampas Cat.” Pampas Cat - Facts, Diet, Habitat & Pictures on Animalia.Bio, 2024). Felines such as jaguars and pumas were portrayed on Moche objects several times, specifically the stirrup spout bottle as these felines were the most powerful predators that lived on the North Coast of Peru, where the Moche were known to reside.

The pampas cats are known to have markings on their fur like domestic cats with colors such as gray-brown, silvery gray and grayish-yellow. However, this stirrup spout bottle depicts the cat in a monochrome manner which highlights the shape and features of its head and body. This monochrome black color was likely achieved by using reduction firing. Reduction firing is the process of firing clay in a kiln that is walled up so that oxygen cannot enter. This causes the iron oxides in the clay to create a black hue. However, this method is not proven to have been used by the Moche; it is simply a hypothesis of what they could have used at the time (Reduction Firing Ceramics.” CUKRASATA,, 2024).

According to the Metropolitan Museum of Art, slip painting was used either for applying to the surface of the pottery before it was fired or it was used to join separate pieces of the clay together. The cat stirrup spout bottle exhibits some brown areas, which may have either faded during excavation, or could be a result of decomposition since its production.

Since these felines are known to be strong and powerful predators of the North Coast of Peru, more often than not the Moche vessels depict them as aggressive and they’re usually attacking prey. In another stirrup spout vessel found in the collections of moche stirrup spout bottles, the cat is depicted ferociously. It has fangs and large claws that are digging into its prey,

the snake (Fig. 11). The Metropolitan Museum of Art's vessel, however, is different, as the cat seems remarkably peaceful ("Stirrup Spout Bottle with Cat: Moche." The Metropolitan Museum of Art, 2024).

The cat seems to be in a resting position with its legs folded underneath it and tail curled on its side. This position is extremely common amongst cats and is also known as the "loaf" position. As per research done by veterinarians, cats sit in this position in order to minimize heat loss and maximize their comfort. Some articles also state that this could be a sign that the cat feels protected. This position allows the cat to both rest and react quickly if necessary (Brown, 2022). The cat's facial expression seems to be neutral, not portraying any specific emotion. Its eyes being open and mouth being closed signals that it's likely relaxed and content. Nonetheless its ears are still up, which means that it is likely to still be alert of its surroundings.

One spout of the vessel is emerging from the head of the cat while the other spout is emerging from the cat's back to join at the top and form a singular spout. This further displays the sculptural excellence by emphasizing the cat's graceful posture. There's a smooth blend between the sculptural representation of the cat and the functional container itself (stirrup spout bottle). This stirrup spout bottle exhibiting the feline truly displays the artistic and technical proficiency of the Moche potters. The item is not only a beautiful piece of art but also a useful artifact, reflecting the Moche culture's aesthetic sense.



Fig. 11

Conclusion

As discussed throughout this paper, it is evident that stirrup spout bottles were not just vessels used by the Moche people for domestic and religious purposes; they are a window into understanding the Moche civilization. As shown in the case study of the Metropolitan Museum's feline stirrup spout bottle, analyzing these vessels provides extremely valuable insights that we might not have been able to get due to lack of literary works. The significance of the stirrup

spout bottle form for study of the Moche civilization demonstrates the importance of examining and analyzing various different objects of material culture when studying a civilization's history.

This analysis also generates new questions for inquiry regarding the stirrup spout vessel. For example, did the elaborate forms of the stirrup spout bottle hinder the practical use of the bottle? If so, how? The paper also raises questions of method, for example what are some of the hindrances archeologists face when trying to preserve these bottles? Finally, a future paper or study might undertake a comparison of the stirrup spout vessel with pottery forms used by other civilizations, prompting the question: what are some similarities between Moche ceramics and other civilizations' ceramics?

Works Cited

- Christopher B. Donnan, "MOCHE CERAMIC TECHNOLOGY," *Ñawpa Pacha: Journal of Andean Archaeology*, no. 3 (1965): 115–124.
- Jules Prown, "Mind in Matter: An Introduction to Material Culture Theory and Method," *Winterthur Portfolio* Vol. 17, No. 1 (Spring, 1982), 3-4.
- "Tangible Things." Harvard University, Harvard Online, 7 Nov. 2023.
- "Moche." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., www.britannica.com/topic/Moche. Accessed 22 July 2024.
- Margaret A. Jackson, "Moche Art" in *Oxford Bibliographies* (2018).
- The urban nucleus is located at the geographical center of the neighborhood similar to how a nucleus is located at the center of the cell.
- Chapdelaine, Claude. *The Growing Power of a Moche Urban Class*. 2001.
- Jeffrey Quilter, "Moche Pottery: Forms, Functions, and Social Change," in *Ñawpa Pacha: Journal of Andean Archaeology*, no. 41, issue 2 (2021): 187-194 and 199-203.
- Huacas Del Sol y de La Luna at 'Moche Site': The Excavation Area of the... | Download Scientific Diagram, www.researchgate.net/figure/Huacas-del-Sol-y-de-la-Luna-at-Moche-site-the-excavation-area-of-the-Franco-peruvian_fig1_228364277. Accessed 29 July 2024.
- Hubert, Erell. "Figuring Identity in Everyday Life." *Journal of Anthropological Archaeology*, vol. 44, Dec. 2016, pp. 1–13, doi:10.1016/j.jaa.2016.07.015.
- Trever, Lisa. "A Moche Riddle in Clay: Object Knowledge and Art Work in Ancient Peru." *The Art Bulletin*, vol. 101, no. 4, 2 Oct. 2019, pp. 18–38, doi:10.1080/00043079.2019.1602449.
- "Classroom Activity: Stirrup Bottle." Virginia Museum of Fine Arts, vmfa.museum/learn-archive/activities/classroom-activity-stirrup-bottle/. Accessed 30 July 2024.
- Slip paint is a liquid mixture of clay and water, which is applied to the surface of the stirrup spout bottle before it is fired. "Stirrup-Spout Bottle: Cupisnique." The Metropolitan Museum of Art, [www.metmuseum.org/art/collection/search/310494#:~:text=Stirrup%2Dspout%20bottles%20\(the%20name,unequalled%20technical%20and%20artistic%20skill](http://www.metmuseum.org/art/collection/search/310494#:~:text=Stirrup%2Dspout%20bottles%20(the%20name,unequalled%20technical%20and%20artistic%20skill). Accessed 31 July 2024.
- "Picture Worlds: Greek, Maya, and Moche Pottery." Getty Museum. Accessed August 24, 2024. https://www.getty.edu/art/exhibitions/picture_worlds/.
- Robinson, Howard. "Dualism." *Stanford Encyclopedia of Philosophy*, September 11, 2020. <https://plato.stanford.edu/entries/dualism/>.
- Koons, Michele L. "To Unite and Divide: Canals, Tinku, Liquids and Time in the Moche World." *Cambridge Archaeological Journal*, vol. 32, no. 4, 17 May 2022, pp. 689–706, doi:10.1017/s0959774322000117.

Horizonte, Prefeitura de Belo. "Pampas Cat." Pampas Cat - Facts, Diet, Habitat & Pictures on Animalia.Bio, animalia.bio/pampas-cat. Accessed 01 Aug. 2024.

Reduction Firing Ceramics." CUKRASATA, www.cukrasata.lv/en-gb/reduction-firing-ceramics/#:~:text=In%20the%20process%20of%20burning,bee%20wax%20and%20vegetable%20oil. Accessed 02 Aug. 2024.

"Stirrup Spout Bottle with Cat: Moche." The Metropolitan Museum of Art, www.metmuseum.org/art/collection/search/309304. Accessed 01 Aug. 2024.

Brown, Jackie. "Cat Loaf (a Cat Tucking Its Paws) ." The Spruce Pets, May 9, 2022. <https://www.thesprucepets.com/why-do-cats-tuck-their-paws-5235285>.

Painting a False Picture: Switzerland's Economic and Military Activity with the Nazis & Allies Overlooking their Declared Neutrality During World War II, 1938-1945

By Wanli Zhang

Abstract

Switzerland's role during World War II, despite its declared neutrality, was characterized by significant economic and military collaboration with both Nazi Germany and the Allies. This paper explores Switzerland's complex involvement in the war, focusing on its financial support to the German military, its strategic positioning for the Axis forces, and the eventual shift towards aiding the Allies. The analysis examines the Swiss banks' substantial financial transactions with Nazi Germany, including the transfer of looted gold and military resources, alongside Switzerland's geographical advantage in facilitating the German war effort. Furthermore, the paper discusses Switzerland's transition from supporting Germany to aligning with the Allies as the war progressed, emphasizing the pivotal role Switzerland's neutral status played in ensuring its survival. Ultimately, the paper argues that Switzerland's deceptive neutrality and economic participation were critical to the outcomes of World War II, influencing both the Axis and Allied powers' military and economic strategies.

Introduction

The Swiss banks, government, and industrialists were deceptive hypocrites during World War II. Switzerland's financial support and economic collaboration with Germany varied at different times during the war. The Swiss profoundly supported the German military from the early to the middle stages, then significantly reduced their financial aid to Germany from the middle stages till the end, and increased economic collaboration with the Allies from 1942 onwards as the end of the German regime became clearer. Switzerland's geographical location in the core of Europe was crucial for the Axis's military and economic success in early and middle stages of the war and the Allies' espionage activities throughout the war. The Swiss government's humanitarian aid to Jewish refugees and the banks' financial collaboration with both parties during the war were essential for Switzerland's survival. The Swiss banks' profound economic support to the German military led to the empire's dominance in earlier stages; however, the Swiss bank's reduction in aid to Germany and increased financial support to the Allies greatly enhanced the Allied military and their authoritative control in Western Europe from the middle stages to the war's end, ultimately leading to an Allied victory, the decline of German supremacy, and the survival of Switzerland.

The Swiss banks' extreme degree of economic aid for the German military and efficiency in transferring assets for the Reichsbank increased the Nazi administration's ability to purchase and produce military resources and support their overseas German-controlled organization, enhancing their military's size and power, ultimately resulting in the German empire's dominance in the early to middle stages of the War. Throughout the War, ledgers determined that the Swiss banks purchased a total of "\$282.9 million worth of Germany's looted gold" which

was shipped to Switzerland by the German Reichsbank from January 1, 1939, to June 30, 1945, though monetary provision was greatly reduced after December 19, 1943 when an Allied-Swiss agreement limited the Swiss Banks' exports to Germany (Swiss DFA Records). The enormous amount of hard currency provided by the Swiss banks allowed the German government "to pay for its war machine" and increase their spending on military supplies such as steel to build Panzer tanks and tungsten to strengthen Messerschmidt fighter planes, enhancing the size and potency of their military and the strength of the German empire (Alan 135; LeBor 30). Additionally, British author Lebor states that "the Swiss National Bank (SNB) was a prime mover in shifting Nazi assets," able to allocate money securely and efficiently to Germany, allowing the German banks to re-transfer and support the Sociedad Financiera Industrial (SOFINDUS), an overseas German-occupied firm in Spain, used to purchase war materials such as ammonium sulfate for explosives and tungsten for weapon-making which were "vital for waging modern warfare." Agents of the Office of Strategic Services (OSS), a U.S. intelligence agency, intercepted that Credit Suisse gave SF700,000 to the Reichsbank to purchase escudos, the Portuguese currency, on 13 October 1944 and that the SNB transferred 100 million USD to Germany which was re-exported to Spain and Portugal (German Gold 240; Allied Claims 240; Lebor 66). The Swiss banks' ability to move assets efficiently to Germany allowed SOFINDUS to receive large sums of money and increase their purchasing power, enhancing the German military's power, ultimately magnifying the empire's supremacy (Lebor 65, 67, 68).

The Role of Switzerland's Geography and Railways in Supporting the Axis Powers

The use of Switzerland's geographical positioning in the center of Europe and the effective lines of transportation led to the success of the German forces in Italy and the growth of their economy due to the ability to move resources quickly and safely, ultimately amplifying Germany's control in Italy during early and middle stages of the war. The German military used the Swiss St. Gotthard line, the quickest transportation route from Germany to Italy, to its full capacity to transport raw materials such as coal, iron, and steel to its forces in Italy (J.R 345). Due to Switzerland's declared military neutrality recognized by the international community in 1815, its railways guaranteed safety and efficiency, permitting German-Italian forces to receive supplies swiftly which helped maintain the size and power of their military. Moreover, the German government was able to reduce the cost for transportation and security because there was no risk of attacks or blockades during transportation, allowing them to allocate their spending for other aspects of the war. Even towards the end of the war when Germany's authority was slowly declining, the railways were still in use to transport resources to Germany and raw materials to Northern Italy to a considerable extent, causing "anxiety" for the Allies (J.R 348). Consequently, these railroads helped maintain the strength of their military and control of Italy under German rule, essentially extending Germany's supremacy in the Italian Republic.

Switzerland's Role in Preventing a German Invasion and Supporting Allied Espionage

Germany's financial dependency on the Swiss banks and its neutral land prevented a German invasion on Switzerland, permitting Allied espionage activities to operate in Switzerland and continue trade between Switzerland and the Allies, enhancing the Allied military's dominance and intel on Nazi affairs and preserving the independence of Switzerland. The German military had plans to invade Switzerland but never proceeded because of their economic reliance on the Swiss banks and their usage of the Swiss railways. Lebor states that "Berlin needed Bern, and a Nazi invasion would have triggered an immediate collapse of the Swiss franc," (168) limiting economic activity for the German government because many countries were unwilling to accept German currency, reducing the Reichsbank's ability to purchase war supplies, diminishing the military's potency (152). Additionally, the tunnels of the St. Gothard and Simplon transportation routes used by the German military were mined with explosives when they were built (J.R 345). The collapse of these lines would cut off Germany's crucial war equipment and communications to Italy and other German-occupied regions, weakening their military's coordination, size, and strength, prohibiting an absolute German rule of Switzerland (LeBor 169). The effective ways in which the Swiss government prevented a German takeover allowed Switzerland to maintain their independence and remain safe during the entire course of the war. Switzerland was able to allow the establishment of the America OSS espionage headquarters on their land which increased the Allied governments' knowledge of the interbank communications and transactions within the financial network linking Germany, Switzerland, and the Iberian Peninsula while ensuring the continuation of economic collaboration with the Allies, enhancing the Allied forces' power (LeBor 64).

Swiss Support to the Allies and the Decline of German Military Power

The Swiss banks and government's unwavering financial support and economic collaboration with the Allies throughout the war created a technology imbalance, enhancing the power and magnitude of the Allied military and the continuity of Switzerland. The Swiss banks maintained a steady flow of trade with the United States throughout the war, providing the U.S. with extensive sums of money of SF132,000,000 for the 341 thousand tons of imported U.S. economic goods and foodstuffs in 1939 and SF151,000,000 for the 213 thousand tons in 1941 (J.R 346, 348). In addition, the United States received a total of 1.178 billion Swiss francs throughout the entire war by exporting gold to Switzerland, indicating that Switzerland was "a major source of hard currency" for the U.S (Cowell 138). As a result, the U.S., a significant ally of the Allied Powers, had a vast reserve of money, allowing the Allies to purchase extensive military supplies, chemicals, technological devices, and precision-instruments such as bomb-timers, causing a technology imbalance, ultimately enhancing the size and capability of the Allied military (LeBor 66; J.R 347). Dr. Max Habicht, a Swiss diplomat of the UN Relief and Rehabilitation Administration, argues that Switzerland's elaborate trade with the U.S. was vital for the "continuation of the Swiss economy" and the preservation of Switzerland's peace because the Swiss people were able to receive valuable goods and adequate nutrition, reducing food

shortages which the country frequently faced, ensuring the health of its population, and enhancing their economy (Habicht 460; J.R 346).

Swiss financial institutions, businessmen, and administration's gradual reduction in monetary provision and economic support for the German power throughout the war decreased the strength of the German military, causing an unsuccessful German post-war operation, eventually leading to the downfall of the German empire and an Allied victory. As the Allies became more potent through Switzerland's financial support, certain Swiss government officials and banks gradually reduced or stopped their economic activities with Germany. Early signs of disengagement occurred in 1940, when I.G. Chemie, a Swiss chemical company, terminated the dividend guarantee arrangement with its German counterpart, I.G. Farbenindustrie, because the Swiss firm wanted "to remove the stigma of the former German interest in the Swiss concern." The separation of the chemical companies, which was a near monopoly, decreased the production of Zyklon B, a lethal poison used by the German military. Two years later, in 1942, the Swiss government denied the manufacturing of military equipment for the German army (J.R 346). In September 1944, the Swiss government banned the export of all military supplies to Germany. These prohibitions greatly reduced the availability of war materials for the German military, decreasing their power while the Swiss increased their financial and military support to the Allies, "meeting all their demands" and enhancing the Allied military's potency. Moreover, Switzerland's new-found policies which were made on February 16 and March 4, 1945, respectively, blocked German credits in Switzerland and prohibited dealings in foreign banknotes (J.R 348). The Swiss government's restrictions due to intense Allied pressure and the government's support for the likelihood of an Allied victory contradicted Germany's post-war operation plans which aimed to establish a secure foundation in Switzerland and place all their funds abroad, hoping the Swiss banks would fully cooperate with (Easton 244). The limitations and reduction in military aid to Germany while increasing the Allied forces' power weakened the German military, eventually leading to the decline of the German regime and an Allied victory overall.

Switzerland's Humanitarian Role and the Jewish Refugee Crisis

Switzerland's gradual increase in the acceptance of Jews throughout the war due to their lengthened support for the Allies caused an increase in recognition and bonding with the Allied powers, enhancing support systems between Switzerland and the Allies which empowered the Swiss economy, obstructed Germany's national goal and preserved the survival of Swiss neutrality. American author Daniel Goldhagen named the Holocaust a 'national project' in which the Swiss government supported from 1938 to August 1942, making Jewish refugees more vulnerable to the German military because of a Swiss-German agreement that German and Austrian Jews' passports would be stamped with a 'J' which led to the refusal of all Jewish refugees seeking asylum in Switzerland (Cowell 142). Consequently, the Swiss government forced them to face the death camps, increasing the number of Jewish deaths and fulfilling Germany's goal of a Jewish massacre. However, the Swiss border became more accessible to

Jewish refugees after June 1942 when the Allies solidified their unity and became more rational in their military and political affairs than the Axis powers, resulting in the loosening of Switzerland's policy on denying Jewish refugees "in parallel with the prospects of an Allied victory." Switzerland accepted 70 thousand Jewish refugees in January 1944, "reached disquieting proportions" by 1943, and was open to Jewish fugitives by the end of the war in 1945 (Stanford; LeBor 222). The Swiss government's contradicting actions impeded Germany's "national project" while turning towards the "Allied point of view." Consequently, the "immeasurable service to humanity" was respected by the Allies, creating a stronger face of legitimacy and forging greater alliances with the Allied forces, increasing trade activities. As a result, Switzerland "scheduled to begin in September 1944" new trade settlements with the Allies to "exchange for wheat, dried fruit, oil, etc." which greatly benefited the Swiss economy and the well-being of its population, allowing Switzerland to prosper and live on as an independent country (J.R 348).

Switzerland's Deceptive Neutrality and Its Impact on WWII

Switzerland's lack of declared neutrality and financial participation in World War II greatly benefited the Nazi Party as well as the Allied forces in their military and economic warfare. Villiger, the Swiss President of 1955, apologized for the poor and inconsiderate decisions "[Switzerland] made at the time in the name of a national interest taken to its narrowest sense," which was essential "for the continuation of Swiss neutrality," showing Switzerland's discredibility and selfish nature. (LeBor 222) Historians are able to understand how Switzerland's drive for survival during World War II forced them to violate their declarations of neutrality which were first made when the French defeated the Swiss confederates at the Battle of Marignano in 1515 and the Treaty of Westphalia of 1648." (Neutrality 1) Politicians from both powers during the war agreed that "Switzerland had a vital right to foreign trade as a neutral during war-time," making scholars ponder the extent of Switzerland's right to engage in economic activities and how their significant participation during World War II at the heart of Europe highlighted their deceptive lies of declared military and economic neutrality as they greatly altered the outcomes of the war (Habicht 459).

Works Cited

- Allied Claims Against Swiss for Return of Looted Gold. 1946.
- Cowell, Alan. Dispatch: Switzerland's Wartime Blood Money. *Foreign Policy*, no. 107, 1997, pp. 132-44, <https://doi.org/10.2307/1149338>.
- Easton, John W., Lt. Col. Letter, *Economic Warfare (Safehaven) Series*: No. 6, 27 Nov. 1944. German Gold Movements (Estimate). Table. 1946, 240.
- Habicht, Max. The Special Position of Switzerland in International Affairs. *International Affairs* (Royal Institute of International Affairs 1944-), vol. 29, no. 4, 1953, pp. 457-63, <https://doi.org/10.2307/2606004>.
- J. R. Switzerland's War-Time and Post-War Policy: Problems of an Encircled Neutral. *Bulletin of International News*, vol. 22, no. 8, 1945, pp. 343-49, <http://www.jstor.org/stable/25643800>.
- LeBor, Adam. *Hitler's Secret Bankers: The Myth of Swiss Neutrality during the Holocaust*. Carol Pub. Group, 1997.
- Neutrality. Last modified 20 Dec. 2021, p.1 <https://www.eda.admin.ch/aboutswitzerland/en/home/politik-geschichte/die-schweiz-und-die-welt/neutralitaet.html#:~:text=Neutrality%20is%20a%20cornerstone%20of,recognised%20Switzerland's%20neutrality%20in%201815>.
- Stanford University News. "Axis Powers Miscalculated After Early Advantages in World War II, Stanford Scholar Says." Last modified 12 Dec. 2017, <https://news.stanford.edu/2017/12/12/axis-powers-miscalculated-early-advantages-wwii-stanford-scholar-says/>.
- Swiss Gold Movements. Table. N.p.: Swiss Official Statement, 1946, 240.
- RG 84: Switzerland State Department and Foreign Affairs Records (The U.S. National Archives and Records Administration), 1, last modified August 15, 2016, <https://www.archives.gov/research/holocaust/finding-aid/civilian/rg-84-switzerland.html#:~:text=On%20December%2019%2C%201943%2C%20after,limit%20their%20exports%20to%20Germany>.
- Weinberg, Gerhard L. "German Plans and Policies Regarding Neutral Nations in World War II with Special Reference to Switzerland." *German Studies Review*, vol. 22, no. 1, 1999, pp. 99-103, <https://doi.org/10.2307/1431584>.

AI Maturity in Fortune 500 Firms: Lessons From Cluster Analysis

By Abhiram Jetty and Shreyas Chittimalla

Abstract

This study investigates the relationship between AI maturity and performance in Fortune Global 500 companies using hierarchical clustering. Analyzing data from the IMD's AI Maturity Index and Fortune Global 500, the study categorizes 192 companies into four distinct clusters based on AI maturity level and key performance indicators (revenues, profits, and employee count). The findings reveal a strong correlation between higher AI maturity and superior financial performance, with companies in higher maturity clusters exhibiting significantly better revenue and profitability. This exploratory analysis offers valuable insights for organizations seeking to enhance their AI capabilities and competitive advantage.

Introduction

Artificial intelligence (AI) has a profound impact on our lives, economy, and world in many different ways (Szczepański 1). In fact, AI could potentially deliver additional economic output of around \$13 trillion by 2030 (Bughin et al. 1). By streamlining operations through the automation of repetitive tasks and derivation of actionable insights from datasets, AI technologies have evolved into essential catalysts for innovation and operationality (Aiden and Michael). As the growth of AI integration intensifies, organizations are encountering mounting pressure to strategically adopt AI capabilities (Eisenberg et al.). Nevertheless, the degree to which organizations are able to effectively utilize artificial intelligence in their decision-making and various processes, commonly termed their "AI maturity," varies widely across organizations in different industries and companies (Reuel et al.). Gaining insights into these disparities is crucial for recognizing optimal practices, addressing obstacles, and enhancing the capacity of AI to achieve quantifiable business results.

The Fortune Global 500 list annually provides a comprehensive list of the top five-hundred companies globally in terms of sales revenue, encompassing a wide range of industries and sectors (Ogrean 136). By virtue of scale, resources, and global presence, these corporations are distinctly placed to champion the adoption and integration of artificial intelligence. However, there is a divide between approaches to AI maturity because of organizational cultures, technological frameworks, strategic goals, and industry practices (McElheran et al.). For some, artificial intelligence promotes innovation and growth, hence supporting substantial revenue growth and customer satisfaction (Jia and Wang). For other entities, it continues to be inadequately utilized, constrained by restricted capabilities or, like other assets, a lack of sufficient financial investment (Ozili). Such discrepancies not only affect the performance of individual companies but also mold overarching trends within industries, rendering the examination of AI maturity among Fortune Global 500 companies a matter of both scholarly and practical importance.

Current scholarly work has predominantly concentrated on theoretical models or case

analyses of individual organizations. Very few empirical studies have tried to examine the adoption patterns of artificial intelligence in a wide and diverse range of firms. The purpose of this study is to address this gap by providing an exploratory analysis of AI maturity in Fortune Global 500 companies, using hierarchical clustering techniques to group firms into distinct classes based on their AI maturity and key performance metrics. Using a dataset of 192 organizations, this research outlines key characteristics of such clusters, highlighting the similarities and differences in their approaches to artificial intelligence.

The objectives of this study are as follows. First, it aims to provide a broad understanding of the differences in AI maturity between Fortune Global 500 firms. Second, it aims to extract actionable insights for organizations looking to enhance their AI capabilities, be it through observations of companies with higher AI-maturity or recognizing common barriers to adoption. Finally, it contributes to the academic debate on how AI is integrated by taking an evidence-based stance on whether AI maturity is related to better organizational performance.

The structure of this paper is organized as follows. The next section discusses existing literature on AI maturity. It will cover the findings, methodology, omissions, and shortcomings of these studies and explains how this study builds on the existing body of research. The methodology section describes the sources of the data as well as analytical methods adopted in the study, with considerable emphasis on hierarchical clustering and its usefulness in bringing out meaningful patterns from the information. The results section describes findings from the research, which details what each cluster exhibits, what these features imply concerning business performance, and what actionable steps can be taken to improve business performance using AI. The paper concludes by discussing the importance of the present study, possible limitations, and future directions for similar studies.

At a time when artificial intelligence is fast becoming a key capability for corporate success, understanding how it is incorporated into leading companies becomes very critical. This study, therefore, explains not just the current state of AI maturity in Fortune Global 500 companies but also provides a strategic framework for using AI as a tool for driving innovation, accelerating growth, and sustaining a long-term competitive advantage for both large corporations and companies newly entering the AI space.

Literature Review

Many articles cover the integration of AI into companies, providing insight into both the initial adoption and advanced integration of these technologies. These articles largely cover topics such as data management, algorithm deployment, and cultural readiness. However, they can often only provide limited insight because they are focused on one industry, a small sample size of companies, or companies with narrow performance metrics. They may also have strict rules on what companies are surveyed, introduce bias through interviews, lack actionable suggestions for companies, and concentrate on theoretical applications. On the other hand, this study uses a large sample size of 192 Fortune Global 500 companies from various industries and clusters them based on business performance metrics and AI maturity instead of industry, which

allows for actionable advice to all companies. Below we review relevant research studies and articles that contribute to the existing body of work on AI maturity.

The research study performed by Sonntag et al. details how AI maturity models are implemented in companies within the manufacturing industry. While it does provide advice to improve maturity models and indicators, the study focuses mainly on maturity models themselves and not the actual AI implementation. Furthermore, it mainly limits itself to the manufacturing industry, doesn't provide specific guidelines for companies based on their specific firm performance variables, and only studies eleven maturity models. On the contrary, this study concerns AI implementation itself and provides targeted advice to companies regardless of their industry.

A study conducted by Dzhusupova et al. provides an extensive framework for Engineering, Procurement and Construction (EPC) companies to integrate AI effectively into the energy sector by covering the entire life cycle of building AI solutions. However, it acknowledges the limitations that its advice may not be applicable outside the energy industry and that its case companies, interviewees, and researchers may introduce bias because of the study's small sample size. This study includes only quantitative data, a relatively large sample size, and draws insights from several industries.

Jankovic and Curovic's research study compares 240 medium and large-sized Serbian companies split into clusters of low, moderate, and high AI adoption. Despite providing significant data and describing how AI technologies enhance data management for such companies, the study fails to provide any specific advice for how these different company clusters could benefit from its conclusions. On the other hand, this study has samples drawn from all over the world and correlates AI maturity with other firm level variables.

An article from Davenport provides significant insight to businesses wanting to use artificial intelligence in their analytical activities. Despite providing actionable advice, it does have its own shortcomings. It mostly covers AI integration for analytics and does not provide much data to back up its findings. It also doesn't specifically target advice to specific clusters of companies. Conversely, this study covers general AI integration, includes data from nearly 200 companies, and provides specific advice for company clusters.

Rahman et al. (2021) covers an extensive review of maturity models based on existing literature. Despite covering maturity models that can apply to any industry, they used 15 maturity models with only limited insight into how they are applied in real businesses. On the other hand, this study studies the real correlation and influence of AI implementation in Fortune Global 500 companies and clusters them to provide actionable advice.

The research article from Durlach et al. covers an AI maturity roadmap, five levels of AI maturity, and how AI can help alleviate the burden of healthcare providers. It does not, however, use any real-world examples of companies or implementations. Despite offering useful insights targeted toward companies in the healthcare industry, it stops short of actionable advice and instead solely creates benchmarks to delineate progress. In contrast, the current research uses hierarchical clustering to study AI maturity of firms from nearly all industries and provides

specific suggestions based on performance levels.

Oleskow-Szlapka et al address AI integration pitfalls and promises in their research paper, with a particular focus on how most companies struggle to implement effective AI. It does not, however, provide targeted advice for specific industries or companies. While it provides valuable information about how to further implement AI technologies into business operations to generate value, it did not survey an even selection of companies: only one retail company, one banking company, and one robotics company. It also suffers from potential interviewer bias. This study, in contrast, utilizes an even selection of Fortune Global 500 companies from which solely quantitative data is used to provide advice to specific company clusters.

The article from Hansen et al. discusses how companies can leverage artificial intelligence to enhance production and warehousing. Specifically, the article also presents methods to implement AI and tools that can be used to address problems with big data and optimization. Despite discussing both the benefits of AI in business like productivity gains and the detriments like widespread unemployment, it does not discuss advice for companies based on their firm-level variables. In contrast to the article, this research study uses hierarchical clustering based on firm-level variables and AI maturity to offer practical suggestions to companies in order to improve overall performance.

Our review of prior work suggests that although past studies have contributed to the study of AI maturity, several drawbacks are evident, such as smaller sample sizes, lack of generalizability, and limited coverage of studies using quantitative data. Our work aims to address these limitations by drawing a larger sample from across the world and by using clustering techniques to create meaningful segments of companies using AI. Additionally, we establish correlations with relevant firm-level variables. We provide our research methodology in the section below.

Research Methodology

Research Methods

We collected data from two secondary sources. The data on our primary variable, AI maturity, was sourced from the research organization IMD (IMD.org, 2025). IMD gathers data on AI maturity by examining three aspects associated with a company: content similarity analysis, the frequency of AI-related keywords, AI deals & investment amounts. For content similarity analysis, IMD uses machine learning algorithms to analyze corporate documents—such as annual reports, analyst calls, and press releases—by comparing them to recognized AI literature. This analysis evaluates both lexical and semantic similarities to determine how well a company's communications align with AI-related content. To measure the frequency of AI-related keywords, IMD employs a curated list of keywords categorized into 16 categories to assess their occurrence in corporate texts. Companies receive scores based on the number of categories referenced, indicating their engagement with various AI concepts. Lastly, to evaluate AI investments and deal activities, IMD relies on data from PitchBook to count and

assess the value of AI-related deals and investments made by each company from 2019 to 2024. This metric reflects the level of financial commitment and strategic investment in AI initiatives. The overall AI maturity score for each company is derived from the total scores in these three areas. Data on control variables such as profits, employee count, and revenue is obtained from Fortune (Fortune.com, 2025). Revenue in this case refers to total income generated by a company from its core operations and includes sales revenue, interest income, and income from joint ventures or subsidiaries. This data is sourced from financial statements, SEC filings, annual reports, and earnings releases. Profit is considered the amount of money a company retains after subtracting all expenses like taxes, interest, and operating costs from total revenue. It is sourced from income statements. The number of employees of a company is the total number of individuals employed by the company at the end of the fiscal year and is sourced from annual reports, employee filings, press releases, and official statements. A part-time employee is counted as half of a full-time employee.

Our sample size of 192 companies comes from the 2024 Fortune Global 500, a list of the 500 largest public or private companies in the world. First, the 500 companies were crosschecked with the IMD AI Maturity Index in order to obtain 192 companies which had both official Fortune data on revenue, profit, and number of employees as well as an AI Maturity rating from IMD. This data is critically important for conducting a robust analysis of the relationships between AI adoption and organizational success.

This data was then compiled in a Microsoft Excel spreadsheet and standardized for revenue, profits, and number of employees to bring all of them to the same scale. Subsequently, all 192 companies were clustered into 4 clusters using the K-means clustering algorithm. Dendrograms were then generated for each cluster and a table showing the cluster means was created.

Analysis Techniques

A clustering method known as hierarchical clustering was employed using JMP Pro software to categorize organizations based on their standardized metrics. This process involves determining the optimal number of clusters (4 was determined to be most effective) and analyzing the groups formed to delineate meaningful distinctions based on AI maturity levels. The choice of four clusters is well supported by multiple validation methods. The canonical plot as shown in Figure 1 from the Discriminant analysis shows reasonable separation among clusters, indicating meaningful distinctions among them. Additionally, the dendrogram analysis confirms that four clusters provide a natural and well-separated grouping of the data. The R^2 values for key variables—AI maturity (0.87), revenue (0.91), profits (0.88), and employees (0.94)—demonstrate that the clustering solution captures a high proportion of the variance, further justifying its robustness. Collectively, these metrics confirm that four clusters provide an optimal balance between interpretability and statistical validity.

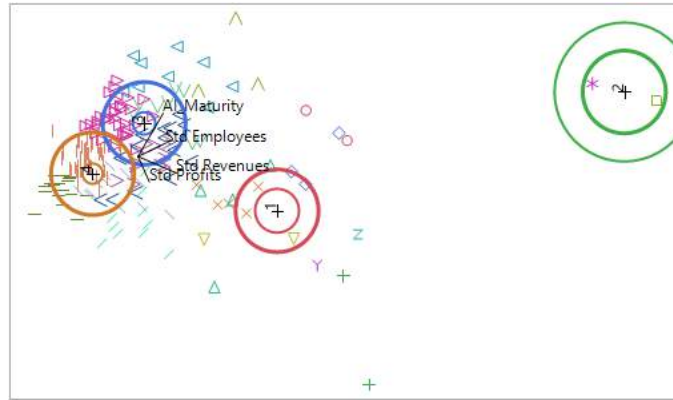


Figure 1. Canonical Plot showing decent separation of clusters

Results of Analysis

We used AI maturity, standardized revenues, standardized profits, and standardized employees as part of our analysis. Figures 2, 3, and 4 show the results in dendrograms and Figure 5 shows the means of the clusters. We present our results below.

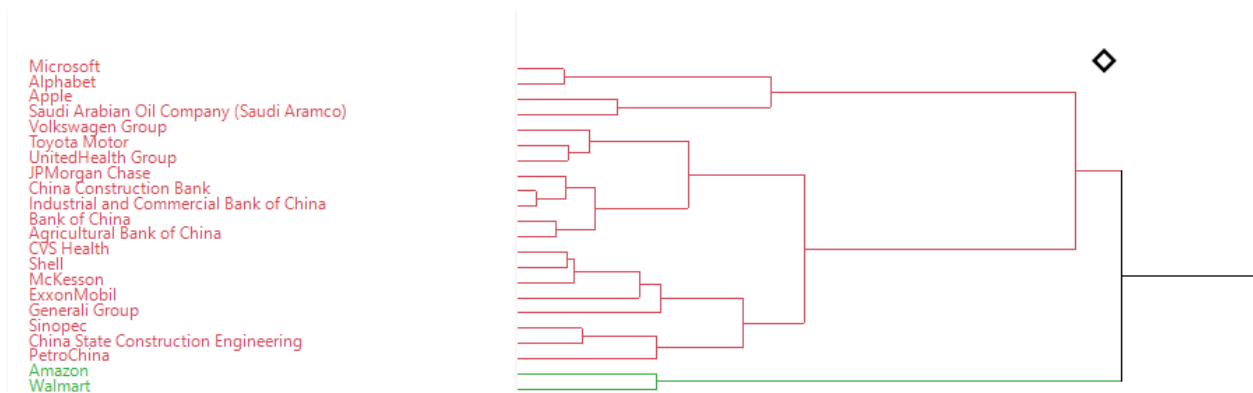


Figure 2. Dendrogram on clusters 1 and 2

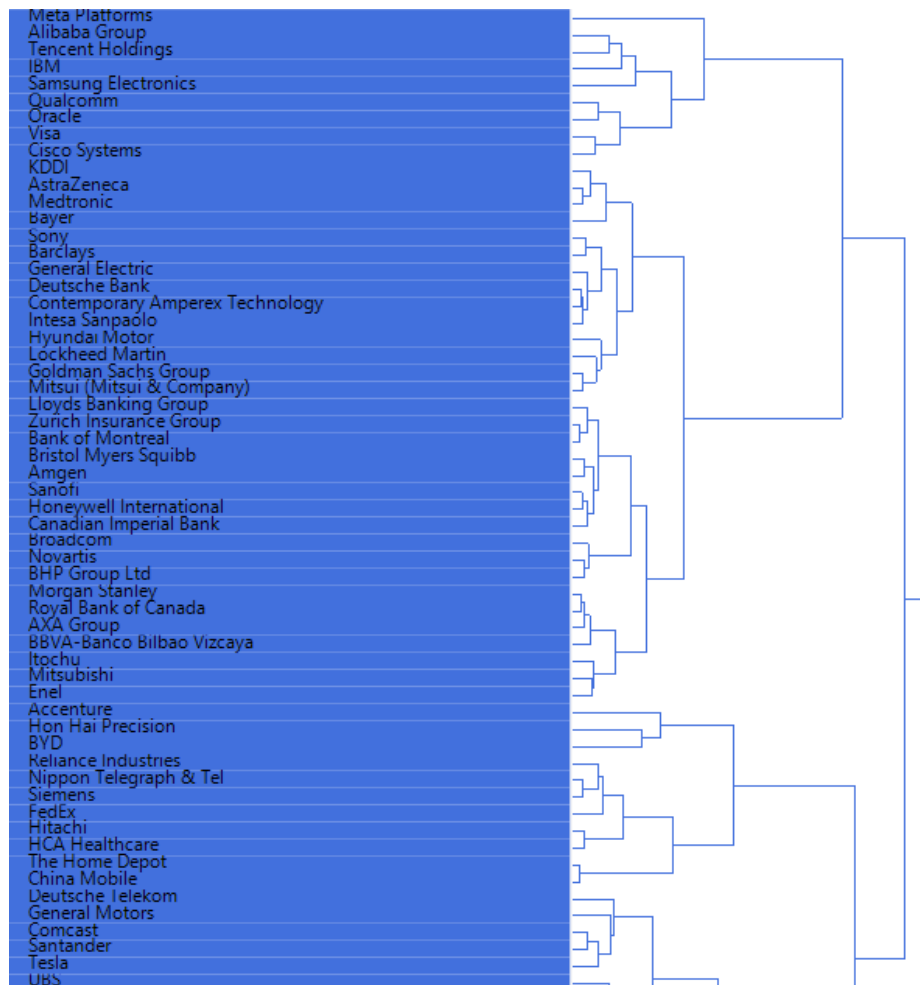


Figure 3. Dendrogram on cluster 3

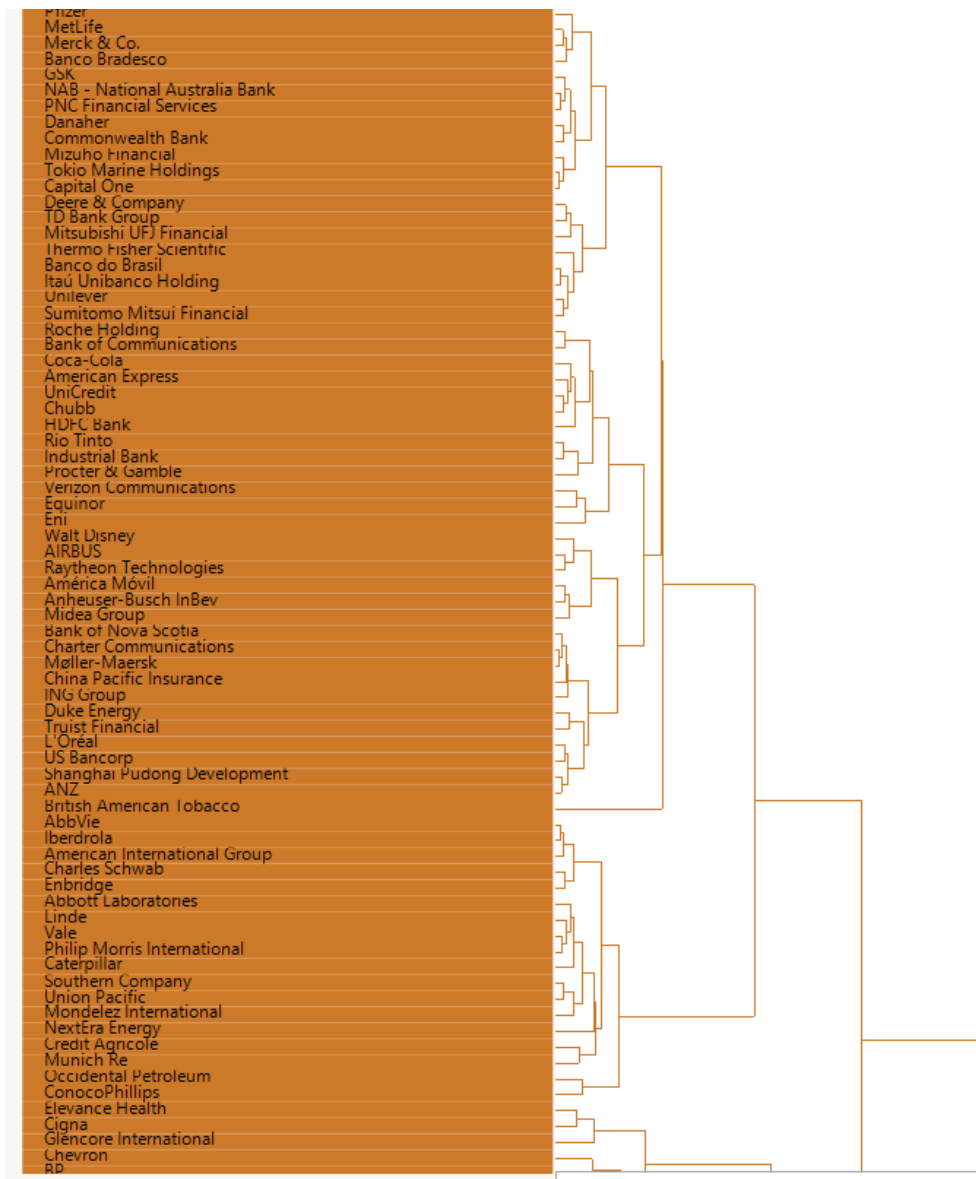


Figure 4. Dendrogram on cluster 4

Cluster Means						
Cluster	Count	AI_Maturity	Std Revenues	Std Profits	Std Employees	
1	20	39.7400	1.9709	1.7559	0.6678	
2	2	66.1000	4.6497	0.7432	7.1861	
3	77	46.0545	-0.1945	-0.0612	-0.0171	
4	93	20.4796	-0.3628	-0.3430	-0.2840	

Figure 5. Cluster means

Cluster 2 had the highest average AI maturity score with the highest means for revenue and employees. It included two companies: Amazon and Walmart. Being retail giants, they also have the highest mean for employee count. Although this cluster has a smaller sample, the results suggest that when organizations have higher AI maturity, they are more likely to use AI into their

business operations to earn more revenue. In contrast, cluster 4 had 93 companies, and they had the lowest average AI maturity score. Correspondingly, their cluster means for standardized revenues and standardized profits were the lowest of all clusters. Many of the cluster 4 companies are banks, financial institutions, or energy companies. Cluster 1 had the second lowest average AI maturity score, but their standardized revenues and standardized profits were superior to that of cluster 3 and cluster 4. This cluster had 20 companies including technology, energy, and automotive companies. Finally, cluster 3 had the second highest average AI maturity score and it included 77 companies. This cluster is largely composed of pharmaceutical, financial, and manufacturing companies. The average revenues and profits for this cluster are below that of cluster 1 and 2.

Results of Analysis

Our results show a stark contrast between companies that have the highest AI maturity levels (cluster 2) and those that have the lowest AI maturity levels (cluster 4). Companies with higher AI maturity levels in general are associated with higher revenue and profit numbers. The groupings within each dendrogram provide additional insights into the similarity of firm performance in terms of AI maturity. For instance, in cluster 3, we see that Meta, Alibaba, Tencent, IBM, Samsung, Qualcomm, Oracle, Visa, and Cisco were grouped together. These technology companies exhibit higher AI maturity levels than Cluster 4 companies like Pfizer, Metlife, Merck & Co, Banco Bradesco, GSK, National Australia Bank, PNC Financials, Danaher, and Commonwealth Bank.

Business Implications

The findings from the cluster analysis offer important business implications for companies aiming to improve their AI maturity. Firms categorized in lower-AI maturity clusters should consider reallocating resources toward AI initiatives, concentrating on impactful areas such as customer service automation, data analytics, and supply chain optimization for quick successes. Developing focused strategies that cater to the specific traits of each cluster can result in more effective outcomes; for example, firms in “tech-savvy” clusters could pursue advanced generative AI applications while those in traditional sectors might benefit from a more gradual integration process. Furthermore, building collaborations with technology providers, startups, and research institutions can expedite the AI journey by providing essential expertise and resources. Investing in AI talent development through training programs and efforts to attract skilled AI professionals is vital for developing necessary capabilities. Companies should also foster a culture that encourages innovation, welcomes experimentation, and seeks leadership involvement in prioritizing AI as a strategic focus. Establishing key performance indicators (KPIs) to track the results of AI projects and creating feedback mechanisms will facilitate ongoing improvement. By recognizing and acting on these implications, companies can effectively elevate their AI maturity and sustain a competitive advantage in an increasingly technology-driven environment.

Conclusion

In conclusion, our work offers valuable insights that can guide organizations in enhancing their AI maturity and overall competitiveness. By identifying distinct clusters based on AI maturity, we show unique characteristics for each group. This nuanced understanding allows for companies who are lagging behind in AI maturity to have more targeted AI investments in high-impact areas, fostering innovation, and developing talent to better leverage AI technologies. However, it is important to acknowledge certain limitations in this study, such as the potential impact of external market factors and the reliance on available data sources, which may not capture all relevant dynamics influencing AI maturity. Additionally, while the study highlights the relationship between AI maturity and organizational strategies, it does not explore extensively the specific challenges faced by industries outside those represented in the analysis. Furthermore, despite providing useful data about the correlation between AI maturity and KPIs, our study does not explore or evaluate the specific methods of AI integration that can produce more desirable outcomes.

Future research should aim to explore these challenges more thoroughly, as well as investigate the long-term impacts of AI maturity on organizational performance across various sectors. Expanding the study to include a larger and more diverse sample could help capture more comprehensive insights into AI adoption trends since this study focuses on just Fortune Global 500 companies. Furthermore, longitudinal studies could provide a clearer picture of how AI maturity evolves over time, allowing for a deeper understanding of the factors driving the successful implementation and integration of AI technologies. Ultimately, by addressing these limitations and expanding upon the initial findings, future research can continue to refine the strategic frameworks necessary for firms to thrive in an increasingly AI-driven landscape.

Works Cited

- “AI Maturity – Methodology.” *IMD Business School for Management and Leadership Courses*, 5 Mar. 2025. Accessed Mar. 17 2025.
imd.org/artificial-intelligence-maturity-index/methodology/.
- “AI Maturity Ranking.” *IMD Business School for Management and Leadership Courses*, 12 Nov. 2024. Accessed Mar. 17 2025.
imd.org/artificial-intelligence-maturity-index/results/#_tab_List.
- Aiden, Dexter, and Lewis Michael. “Artificial Intelligence in Business: Enhancing Operational Efficiency and Navigating Ethical Challenges.” *ResearchGate*, 29 Sept. 2024. Accessed Mar. 17 2025. researchgate.net/publication/384441985.
- Bughin, Jacques, et al. “Modeling the Impact of AI on the World Economy.” *McKinsey Global Institute*, McKinsey & Company, 4 Sept. 2018. Accessed Mar. 17 2025.
mckinsey.com/~media/mckinsey/featured%20insights/artificial%20intelligence/notes%20from%20the%20frontier%20modeling%20the%20impact%20of%20ai%20on%20the%20world%20economy/mgi-notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy-september-2018.ashx.
- Davenport, Thomas H. “From Analytics to Artificial Intelligence.” *Journal of Business Analytics*, Taylor & Francis, 19 Nov. 2018. Accessed Mar. 17 2025.
tandfonline.com/doi/full/10.1080/2573234X.2018.1543535.
- Durlach, Peter, et al. “The AI Maturity Roadmap: A Framework for Effective and Sustainable AI in Health Care.” *The New England Journal of Medicine*, NEJM Group, 28 Mar. 2024. Accessed Mar. 17 2025. ai.nejm.org/doi/full/10.1056/AI-S2400177.
- Dzhupova, Rimma, et al. “Choosing the Right Path for AI Integration in Engineering Companies: A Strategic Guide.” *Journal of Systems and Software*, Elsevier, 24 Dec. 2023. Accessed Mar. 17 2025. sciencedirect.com/science/article/pii/S0164121223003400.
- Eisenberg, Ian W., et al. “The Unified Control Framework: Establishing a Common Foundation for Enterprise AI Governance, Risk Management and Regulatory Compliance.” *arXiv*, 7 Mar. 2025. Accessed Mar. 17 2025. arxiv.org/html/2503.05937v1.
- “Fortune Global 500.” *Fortune*, 4 Aug. 2024. Accessed Mar. 17 2025.
fortune.com/ranking/global500/.
- Hansen, Hans Fredrik, et al. “Understanding Artificial Intelligence Diffusion Through an AI Capability Maturity Model.” *SpringerLink*, Springer US, 23 Aug. 2024. Accessed Mar. 17 2025. link.springer.com/article/10.1007/s10796-024-10528-4.
- Jankovic, Svetozar D., and Dejan M. Curovic. “Strategic Integration of Artificial Intelligence for Sustainable Businesses: Implications for Data Management and Human User Engagement in the Digital Era.” *MDPI*, 24 Oct. 2023. Accessed Mar. 17 2025.
mdpi.com/2071-1050/15/21/15208.
- Jia, Yanhuai, and Zheng Wang. “Application of Artificial Intelligence Based on the Fuzzy Control Algorithm in Enterprise Innovation.” *Heliyon*, Cell Press, 18 Mar. 2024. Accessed Mar. 17 2025. [cell.com/heliyon/fulltext/S2405-8440\(24\)04147-1](https://cell.com/heliyon/fulltext/S2405-8440(24)04147-1).

- McElheran, Kristina, et al. "AI Adoption in America: Who, What, and Where." *Journal of Economics & Management Strategy*, Wiley Online Library, 24 Jan. 2024. Accessed Mar. 17 2025. onlinelibrary.wiley.com/doi/full/10.1111/jems.12576.
- "Methodology for Global 500." *Fortune*, 8 Aug. 2023. Accessed Mar. 17 2025. fortune.com/franchise-list-page/global-500-methodology-2023/.
- Ogrean, Claudia. "A Snapshot of the World of Global Multinationals – an Industry Based Analysis of Fortune Global 500 Companies." *Studies in Business and Economics*, 2017. Accessed Mar. 17 2025. researchgate.net/publication/320588763_A_Snapshot_of_the_World_of_Global_Multinationals_-_An_Industry_Based_Analysis_of_Fortune_Global_500_Companies.
- Ozili, Peterson K. "Impact of Digital Finance on Financial Inclusion and Stability." *Borsa Istanbul Review*, Elsevier, Dec. 2018. Accessed Mar. 17 2025. sciencedirect.com/science/article/pii/S2214845017301503.
- Rahman, Adul Hadi Abd. "Artificial Intelligence Maturity Model: A Systematic Literature Review." Edited by Raghad Baker Sadiq et al., *PeerJ Computer Science*, PeerJ, 25 Aug. 2021. Accessed Mar. 17 2025. peerj.com/articles/cs-661/.
- Reuel, Anka, et al. "Responsible AI in the Global Context: Maturity Model and Survey." *arXiv.Org*, 13 Oct. 2024, arxiv.org/abs/2410.09985.
- Sonntag, Martin, et al. "Development and Evaluation of a Maturity Model for AI Deployment Capability of Manufacturing Companies." *Information Systems Management*, Taylor & Francis, 29 Feb. 2024. Accessed Mar. 17 2025. tandfonline.com/doi/full/10.1080/10580530.2024.2319041.
- Szczepański, Marcin. "Economic Impacts of Artificial Intelligence (AI)." *European Parliamentary Research Service*, European Parliamentary, July 2019. Accessed Mar. 17 2025. [europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI\(2019\)637967_EN.pdf](https://europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf).
- Oleskow-Szlapka, Joanna et al. "Striving for Excellence in AI Implementation: AI Maturity Model Framework and Preliminary Research Results." *Scientific Journal of Logistics*, Nov. 2019. Accessed Mar. 17 2025. researchgate.net/publication/337561922_Striving_for_Excellence_in_AI_Implementation_AI_Maturity_Model_Framework_and_Preliminary_Research_Results.

Structural Analysis of Virulence Factors Between *S. aureus* Strains MSSA476 and MRSA252 By Linda Yu

Abstract

Staphylococcus aureus is a common bacterium that typically does not cause harm but may cause severe and fatal infections. *S. aureus* bacteria can be classified into two types: MSSA (methicillin-sensitive *S. aureus*), strains susceptible to antibiotics, and MRSA (methicillin-resistant *S. aureus*), strains resistant to methicillin-based antibiotics. Incidence rates of *S. aureus* infection differ between geographical locations, with incidence rates in the U.S. at 0.05% of the population. Despite its low incidence rates, infections can cause death, with a mortality rate of approximately 13%. My objective was to conduct a comprehensive literature review and computationally simulate interactions between virulence factors of *S. aureus* and human immune system factors. To achieve this, I analyzed differences in the genetic makeups of two *S. aureus* strains, MSSA476 and MRSA252, to identify specific genes linked to virulence factors and infection. Structural analysis reveals that MSSA476 and MRSA252 have differences in the manner in which they interact with human PsGL1 proteins and proteins such as the SAG superfamily.

Background

Staphylococcus aureus is a gram-positive bacterium that can be commonly found on the skin of individuals without causing any harm. However, if the bacterium is allowed to enter the bloodstream, it can cause severe, life-threatening infections (CDC). These infections primarily affect immunocompromised individuals or those with pre-existing health conditions. Other people who are more likely to be affected are those who work in healthcare. Incidence rates of *S. aureus* infection differ between geographical locations, with incidence rates in the U.S. at 0.05% of the population—approximately 167,000 people annually (J van Hal et al.). Despite its low incidence rates, infections can become severe and cause death, with a mortality rate of approximately 13%. *S. aureus* is classified as MSSA (methicillin-sensitive) and MRSA (methicillin-resistant), with the latter posing a greater treatment challenge. Because *S. aureus* bacteria of the MRSA strain are resistant to methicillin-based antibiotics, they can be harder to treat using traditional methods, so patients infected have higher risks of mortality than those infected with bacteria of the MSSA strain (Adeiza and Aminul).

Beyond antibiotic treatment susceptibility, an additional danger of *S. aureus* infections includes its ability to cause the life-threatening condition of toxic shock syndrome (TSS). Often, TSS infections are caused by toxins produced by staph bacteria, but the condition can also be caused by group A streptococcus bacteria (Mayo Clinic). Some risk factors for TSS include using tampons, having surgical wounds, and having a viral infection. However, it is important to note that anyone can be affected by TSS. TSS presents with flu-like symptoms, including sudden high fever, a sunburn-like rash, and muscle aches. If the infection spreads, accumulating toxins in the blood may lead to shock. During TSS, superantigens from *S. aureus* result in the

overactivation of cytokines and inflammatory cells, such as neutrophils (Ross and Shoff). This extreme activation of immune cells is a component of toxic shock syndrome (Xu et al.).

In this paper, I will focus on Staphylococcal Superantigen-like protein 11 (SSL11) and the role it may play in causing toxic shock syndrome. In previous studies, it was shown that SSL11 from a strain of MRSA arrested cell motility by inducing cell adhesion (C. Chen et al.). Through inducing cell adhesion, SSL11 prevents neutrophil recruitment, which could allow the bacteria to successfully proliferate and release toxins into the bloodstream, triggering TSS.

To further investigate the way these two strains of *S. aureus* interact differently with immune cells, I will also focus on interactions between the two strains with PsGL1 proteins. P-selectin glycoprotein ligand-1 is a glycoprotein that is found on the surface of white blood cells and plays a role in the body's immune response. Specifically, PsGL1 on leukocytes has proven to have an important role against invasive pneumococcal disease (Ramos-Sevillano et al.). The interactions between human PsGL1 proteins and MSSA476 and MRSA252 could show how the two bacteria differ in the way they invade and affect human hosts. Further, I will also model the interactions between PsGL1 proteins from mice and these two strains of *S. aureus* to see if there are further differences, as mice are unable to experience septic shock the same way humans do (P. Chen et al.).

The purpose of this paper is to analyze the differences between two strains of *S. aureus*, MSSA476 and MRSA252, and identify the genetic differences between the two that are linked to differing virulence factors. Along with modeling the way these two strains interact with the human immune system, this paper could provide insight into the different ways MSSA and MRSA bacteria behave and contribute to future research that could treat MRSA infections better and decrease the difference in mortality rates between the two strains.

Methods

Creating and analyzing a Synmap

I used genomevolution.org to create a synmap to compare the genomes between the *S. aureus* strains MRSA252 and MSSA476. Afterward, to analyze one specific difference I identified in the Synmap, I used the 'Compare in Gevo' feature of this tool. Along with this tool, I used VFDB BLAST to identify the specific genetic difference between the two strains at the genomic region of interest identified within the Synmap (See Fig. 1).

Multi-protein alignment

I used Clustal Omega to conduct a multi-protein alignment to generate data regarding the specific protein differences between SSL11 in each strain. Use the FASTA sequences in the Appendices.

Modeling structures and interactions of proteins

Finally, to compare the 3D structures of the SSL11 of each strain in more detail, I used RCSB PDB - Pairwise Structure Alignment Tool to model them together using the existing AlphaFold structure predictions. (RCSB_MRSA252_MSSA476_3D) To model the interactions between the SSL11 of each strain and the P selectin ligands of both species, I used AlphaFold Server.

Families and domains

To identify protein structural components such as families and domains of SSL11 from MRSA252 and MSSA476, I used InterProScan - InterPro.

Results

S. aureus strains MSSA476 and MRSA252 are highly conserved with few unique genetic shifts, including protein SSL11

I chose the two strains, MSSA476 and MRSA252, mainly to determine the difference between the fact that one is methicillin-susceptible and the other is resistant, and each may contribute differently to the development of TSS. Ultimately, these small and large genetic differences and similarities might heavily affect the way both infections are treated and their likelihood of producing TSS.

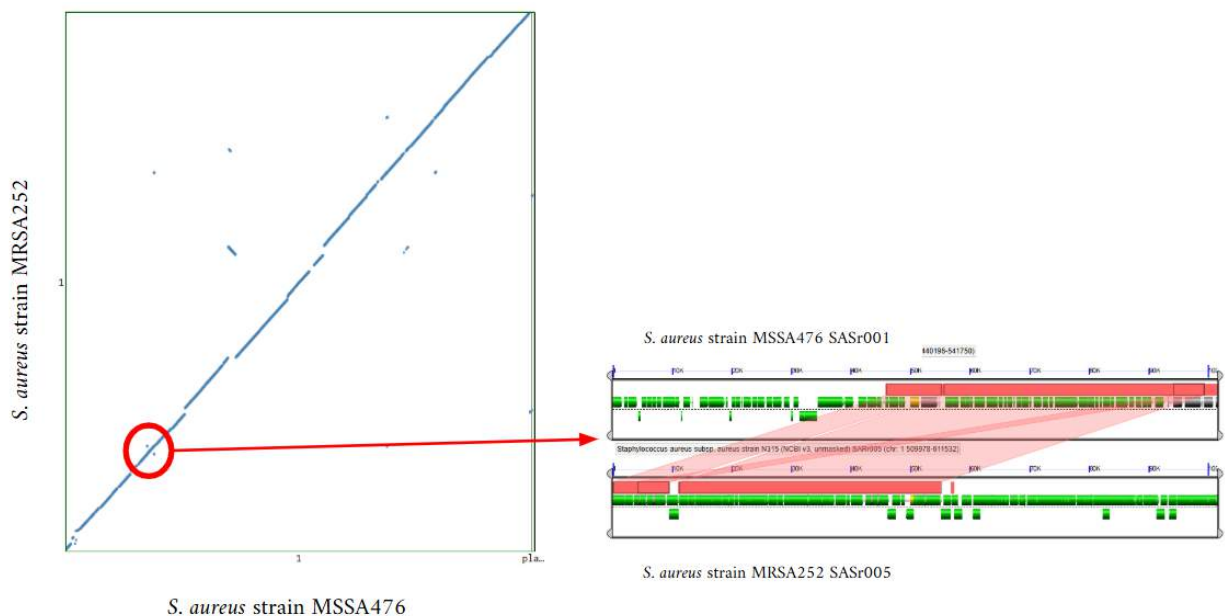


Fig 1: Using genomevolution.org, I created a synmap to identify differences between the genomes of *S. aureus* strain MRSA252 and *S. aureus* strain MSSA476. Two of the differences that I identified were located directly diagonally along the line, and I chose to do further analysis on these two points. Using the Compare in Gevo feature, I analyzed one of the differences in MSSA476 SASr001 and MRSA252 SASr005, and found that there were some similarities shown in red with many differences. After, using VFDB, I identified that one of the major differences between MSSA476 SASr001 and MRSA252 SASr005 was the gene SAS0396 found in MSSA476 and not in MRSA252. This gene codes for the superantigen-like protein SSL11.

I created a Synmap using genomevolution.org to identify differences between the genomes of *S. aureus* strain MRSA252 and *S. aureus* strain MSSA476. Two of the differences that I identified were located directly diagonally along the line, and I chose to do further analysis on these two points. Using the Compare in Gevo feature, I analyzed one of the differences in MSSA476 SASr001 and MRSA252 SASr005 and found that there were some similarities shown in red with many differences. Using VFDB, I identified that one of the major differences between MSSA476 SASr001 and MRSA252 SASr005 was the gene SAS0396 found in MSSA476 and not in MRSA252. This gene codes for the superantigen-like protein SSL11.

Researching further into this region of interest, I identified that this region contained a major difference between the two sequences of the strains in this region. I found that a gene, SAS0396, which codes for the superantigen-like protein SSL11, could be found in the sequence of MSSA476 in this region, while it could not be found in the sequence of MRSA252 in this region.

The role of SSL11 and its structural differences in MSSA476 and MRSA252

SSL11 is a superantigen-like protein that can be found in strains of *Staphylococcus aureus*. Previous research has found that bacteria can arrest cell motility in dHL60 cells—cells used to model neutrophils—by inducing cell adhesion through multiple domains of SSL11 (Yang et al.). It can be inferred that this protein plays a role in the behavior of *S. aureus* and the way it infects the human body; however, previous research has not indicated whether it affects the bacteria's ability to resist methicillin.

To show this, I compared the structures of the SSL11 from each strain of *S. aureus*. Figure 2 shows a screen capture of the 3D alignment of the two protein structures based on their AlphaFold-generated structure. As the figure shows, the structures of the two proteins overlap consistently, suggesting that the functions of most of the two proteins are similar. However, there are specific regions where the structures differ quite a bit. There are many gaps between the two protein structures in a few regions, suggesting that the proteins may differ in function in these regions.



Fig 2: Using <https://www.rcsb.org/alignment>, I did a 3D comparison between the SSL11 proteins of MRSA252 and MSSA476. In the model, the orange highlighted protein is the SSL11 of MRSA252, while the blue highlighted protein is the SSL11 of MSSA476. Looking at the model, it is clear that the protein structure of both strains of *S. aureus* is very similar, with many parts of their structures directly overlapping in the model. However, other parts of the structure are not overlapping, as can be seen in the top left corner of the model, where the proteins have a large gap in between them, indicating a difference in the amino acid sequence of the protein in that area, and a potential difference in function.

To further compare the structural differences between the SSL11 of MSSA476 and MRSA252 and explore the role of the structural differences in functional changes in immune response outside of methicillin resistance, I used the AlphaFold server to model their interactions with P-selectin glycoprotein ligand-1 from humans and mice. Figure 3 shows the models of interactions between the SSL11 proteins of each strain and the glycoproteins of each species.

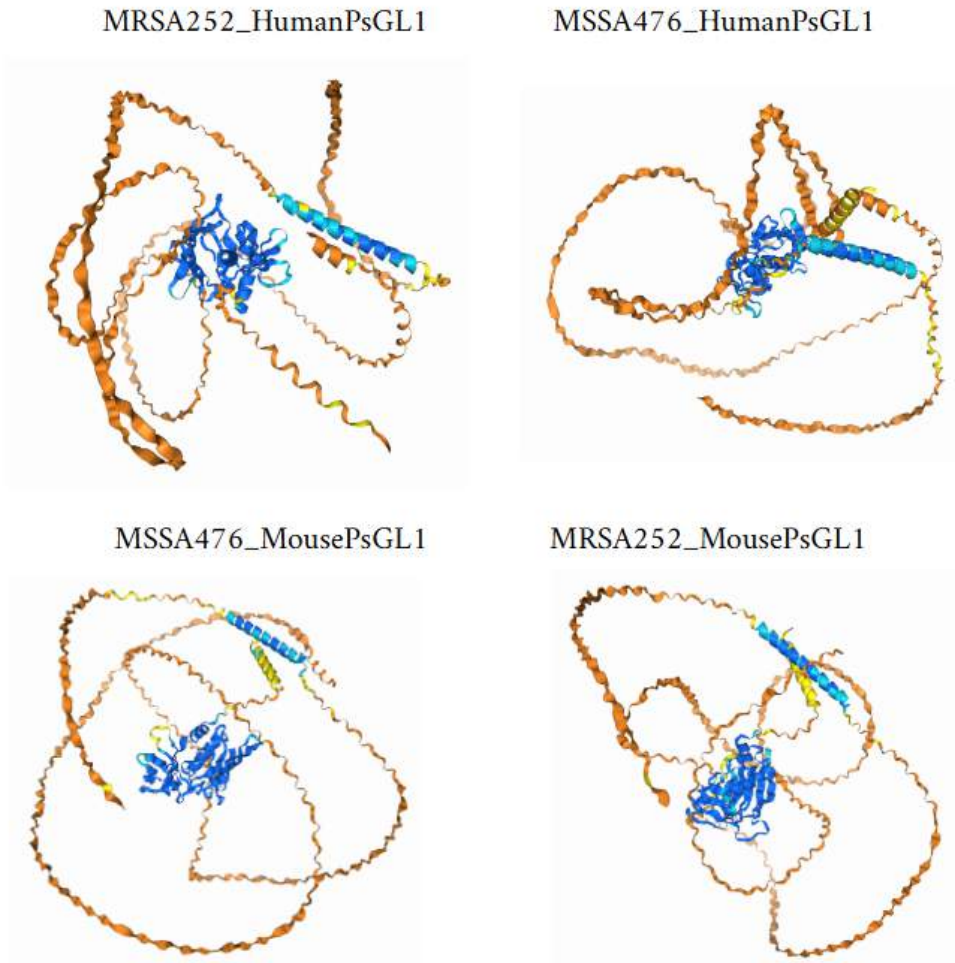


Fig 3: Using <https://alphafoldserver.com/>, I modeled the interactions between two strains of *S. aureus*, MRSA252 and MSSA476, and the P-selectin glycoprotein ligand-1 of humans and mice. The dark blue in the figures represent areas of higher confidence, while the orange represents areas of lower confidence. The models appear to show that MSSA476 interacts with Human PsGL1 in three areas, while MRSA interacts with Human PsGL1 in only two areas. Further, MSSA476 interacts with Mice PsGL1 in two areas, while MRSA interacts with PsGL1 in only three areas.

As shown in the model, the SSL11 proteins from each strain interact differently with the PsGL1 glycoproteins from each species. They interact in different areas, suggesting that the interactions differ in function. The differences in interaction between the SSL11 proteins and the human and mouse PsGL1 glycoproteins could suggest that humans and mice have different immune responses toward *S. aureus* infection. The differences in interaction between the two SSL11 proteins and the human PsGL1 could account for the differences in the way they must be treated. These different interactions could affect how one strain, MSSA476, can be treated effectively with methicillin, while the other strain, MRSA252, cannot be treated effectively with methicillin.

Differences in the function of the SSL11 protein in MSSA476 and MRSA252

While structural differences may imply functional differences, I decided to do more research into the functional differences between the two proteins. As shown in Figure 4, I used an Interproscan to compare the families and domains of SSL11 from MSSA476 and MRSA252. While the sequences had mostly the same families and domains, there were a few regions identified as functionally unique that stood out.

For example, in the top right corner, the SAG superfamily is boxed because it appears in MSSA476 in a much longer sequence than it does in MRSA252. The SAG superfamily is a superfamily of exotoxins that are associated with toxic shock syndrome. This being expressed in a longer sequence in MSSA476 could suggest that this strain produces a larger amount of this exotoxin and may be more likely to result in and may lead to more severe, or faster onset toxic shock syndrome. Further, as shown in the figure, MRSA252 contains the C-terminal domain, which MSSA476 does not. This would suggest that the two SSL11 proteins function in different manners as the two bacteria may perform processes such as transcription, RNA splicing, and capping differently.



Fig. 4: Using <https://www.ebi.ac.uk/interpro/search/sequence/>, I compared the families and domains of SSL11 from MRSA252 and MSSA476. Highlighted in the figure are the differences in length of sequences between the two proteins. Also, in red, are the Bacterial superantigen superfamily signature present in MSSA476 and the Superantigen toxin, C-terminal present in MRSA252.

The results found above back conclusions made in previous literature, as there is a difference in the structure and functions of SSL11 in MSSA476 and MRSA252, suggesting that the two strains of *Staphylococcus aureus* have differences in function. The two strains also have differences in proteins, such as the SAG superfamily and C-terminal domain, suggesting that the onset of TSS may be different for infections by the two strains. This would support previous literature and cases, as the two bacteria lead to infections that require different treatments and may benefit from unique immune support related to immune cell motility.

Discussion

This study investigated the role of SSL11 in the different behaviors of *S. aureus* strains MSSA476 and MRSA252. I found that these two strains are remarkably similar, including the fact that both strains contain the protein SSL11. However, the structures, families, and domains of SSL11 from MSSA476 and MRSA252 are different, suggesting that the functions of the two proteins are different as well. Specifically, MSSA476 expresses a longer sequence of the SAG superfamily gene. As SSL11 may primarily work to prevent neutrophil recruitment, this added aspect may suggest that MSSA476 may produce a more intense toxic shock syndrome than MRSA252.

Since MSSA476 is susceptible to methicillin and other antibiotic treatments, this strain having a longer sequence of the SAG superfamily gene could suggest that it produces more exotoxins and produces them faster than MRSA252. As bacteria usually do not keep unnecessary DNA, MSSA476 may have evolved to release exotoxins quicker to severely damage its host before being killed off. MRSA252, on the other hand, may produce fewer exotoxins over a more extended period because it does not face the same pressure as MSSA476 to damage its host as quickly as possible.

Additionally, I reviewed species-specific interactions between SSL11 and PsGL1, as mice are commonly used as models for human diseases. However, mice do not experience the same toxic shock syndrome that humans do, so research studies that use mice as models must take extra steps to create an accurate picture. One specific difference in the ways mice and humans experience TSS is that TSST-1 is species-specific and does not affect mice. This study showed how PsGL1 from mice interact differently with MRSA252 and MSSA476 compared to PsGL1 from humans. This shows the importance of humanizing mice models. Further research should be done to understand the specific differences between the immune cells of humans and mice and the way they interact with *S. aureus* to improve the accuracy of mice models in the research of TSS.

Finally, the current treatment for toxic shock syndrome is antibiotics, whether from MSSA or MRSA infection. However, as we have realized in recent years and as can be seen by MRSA strains, antibiotic resistance has become a large problem (“Antimicrobial Resistance”). Thus, other options to aid in treatment for infections need to be explored to decrease the amount of antibiotics necessary for treating toxic shock syndrome. Since SSL11 prevents neutrophil recruitment, drugs can be given to increase neutrophil production in the body so that neutrophils are more likely to enter the site of infection. Further, as this paper has shown that SSL11 in MSSA476 contains the SAG superfamily in a longer sequence than SSL11 in MRSA252, intravenous immunoglobulin should be administered more in MSSA infections to treat exotoxins produced by the bacterium (Yanagisawa et al.).

In the future, further research should be done to validate that neutrophils are affected differently by these strains of bacteria. To do this, a stronger connection between neutrophils' role in toxic shock syndrome should be established. A meta-analysis can be performed using strains isolated from patients with toxic shock syndrome with varied severity and directly testing

these strains against neutrophils in cell culture to judge their reaction. If there is a difference in how the neutrophils are affected by the strains, it may be influenced by differences in the SSL11 proteins of each strain.

This paper shows that while the sequences of MSSA476 and MRSA252 are remarkably similar, there are a few differences that could account for different functions. Modeling of the two strains overlaid on top of each other shows that the large genetic similarities accounts for large structural similarities, although the structural differences reflect differences in the way the two strains interact with immune cells as well as the features that allow them to produce toxic shock syndrome.

Works Cited

- Adeiza, Shuaibu Suleiman, and Islam Aminul. Meta-Meta-Analysis of the Mortality Risk Associated with MRSA Compared to MSSA Bacteraemia - PMC. 2024, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11142415/>.
- “Antimicrobial Resistance.” World Health Organization, 2023, <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>.
- CDC. “Staphylococcus Aureus Basics.” Staphylococcus Aureus, 15 Apr. 2024, <https://www.cdc.gov/staphylococcus-aureus/about/index.html>.
- Chen, Chen, et al. Staphylococcal Superantigen-like Protein 11 Mediates Neutrophil Adhesion and Motility Arrest, a Unique Bacterial Toxin Action | Scientific Reports. 2019, <https://www.nature.com/articles/s41598-019-40817-x>.
- Chen, Peter, et al. “Differences Between Murine and Human Sepsis.” The Surgical Clinics of North America, vol. 94, no. 6, Dec. 2014, pp. 1135–49. PubMed Central, <https://doi.org/10.1016/j.suc.2014.08.001>.
- J van Hal, Sebastian, et al. Predictors of Mortality in Staphylococcus Aureus Bacteremia - PMC. 2012, <https://pmc.ncbi.nlm.nih.gov/articles/PMC3346297/>.
- Mayo Clinic. “Toxic Shock Syndrome: A Rare but Potentially Fatal Condition-Toxic Shock Syndrome - Symptoms & Causes.” Mayo Clinic, 2022, <https://www.mayoclinic.org/diseases-conditions/toxic-shock-syndrome/symptoms-causes/syc-20355384>.
- Ramos-Sevillano, Elisa, et al. “PSGL-1 on Leukocytes Is a Critical Component of the Host Immune Response against Invasive Pneumococcal Disease.” PLoS Pathogens, vol. 12, no. 3, Mar. 2016, p. e1005500. PubMed Central, <https://doi.org/10.1371/journal.ppat.1005500>.
- Ross, Adam, and Hugh W. Shoff. Toxic Shock Syndrome - StatPearls - NCBI Bookshelf. 2023, <https://www.ncbi.nlm.nih.gov/books/NBK459345/>.
- Xu, Stacy X., et al. Superantigens Subvert the Neutrophil Response To Promote Abscess Formation and Enhance Staphylococcus Aureus Survival In Vivo - PMC. 2014, <https://pmc.ncbi.nlm.nih.gov/articles/PMC4187807/>.
- Yanagisawa, Chie, et al. “Neutralization of Staphylococcal Exotoxins in Vitro by Human-Origin Intravenous Immunoglobulin.” Journal of Infection and Chemotherapy: Official Journal of the Japan Society of Chemotherapy, vol. 13, no. 6, Dec. 2007, pp. 368–72. PubMed, <https://doi.org/10.1007/s10156-007-0551-6>.
- Yang, Chen, et al. “Multiple Domains of Staphylococcal Superantigen-like Protein 11 (SSL11) Contribute to Neutrophil Inhibition.” Biochemistry, vol. 61, no. 7, Apr. 2022, pp. 616–24. PubMed Central, <https://doi.org/10.1021/acs.biochem.2c00018>.

The Application of the SIR Framework to Analyze COVID-19 Transmission Patterns

By Nipun Saxena

Abstract

The COVID-19 pandemic has shown us the need for efficient public health measures and a better understanding of how diseases spread. This study applies the Susceptible-Infectious-Recovered (SIR) model to analyze transmission trends across six countries (the United States, Canada, India, Russia, Brazil, and China), primarily focusing on variations in infection (β) and recovery (γ) rates. Using weekly COVID-19 data from Our World Data, we segmented the pandemic into distinct waves from countries' local maxima and minima. We then estimated SIR parameters to measure the impact of population density, healthcare capacity, and interventions such as lockdowns and mask mandates, and vaccination campaigns. Our analysis shows that densely populated regions with limited healthcare resources expressed higher β values, while γ rates significantly increased following vaccination efforts. Countries with rapid healthcare, such as Canada and China, demonstrated lower β values and higher γ values, which correlate with stable reproduction numbers (R_0) over time. Strict lockdowns directly correlate with reduced transmission rates, and the role of vaccination programs greatly enhances the recovery rates. Regional differences and data reporting limitations illustrate the need for context-specific public health policies. This study shows us the use of the SIR model in planning for the next pandemic.

Keywords

SIR Model, Covid-19, Epidemiological Parameters, Epidemic Modelling, Disease Transmission

Introduction

“Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus” (World Health Organization, n.d.). The pandemic had an unprecedented impact on global health, politics, societies, and economies. The coronavirus was first detected in China, in late 2019. From then the virus rapidly spread worldwide, throughout 2020-2023. The virus led to over 700 million confirmed cases and several million deaths as of 2024 (COVID-19 Epidemiological Update, n.d.). The COVID-19 pandemic has shown us the importance of understanding infectious disease dynamics to optimize public health responses. This study investigates the variation in the parameters of the Susceptible-Infectious-Recovered (SIR) model (Harko et al., 2014). This study investigates the infection rate (β) and the recovery rate (γ) across countries and regions using Python. Using COVID-19 data, we examine how these parameters are influenced by factors such as public health measures and new variants of the virus. Our research aims to quantify the impact of specific interventions, including vaccination programs, quarantine measures, and mask mandates, on β and γ values. The findings will help us provide insight into which strategy is the most efficient. Hypotheses suggest that β values will be higher

in densely populated regions and countries with limited healthcare resources, while γ values will increase the most by vaccination programs. By analyzing countries' variability in SIR parameters and identifying the most effective interventions, this study contributes to the development of evidence-based, equitable public health protocols. These findings will help policymakers in making the most efficient strategies to mitigate the impact of a future pandemic. Public health measures to control the outbreak, including lockdowns, mask mandates, vaccination campaigns, and travel restrictions. The effectiveness of these interventions varied significantly between countries, and regions, and was heavily influenced by factors such as healthcare capacity, population density, compliance with public health measures, and socio-political conditions.

In this study we will use a mathematical model such as the Susceptible-Infectious-Recovered (SIR) (Prodanov, 2022) model to analyze and predict the spread of infectious diseases, focusing on COVID-19 data. The SIR model simplifies complex epidemiological processes into three key parameters: the infection rate (β), the recovery rate (γ), and the basic reproduction number (R_0), which represents the average number of secondary infections caused by a single infected individual. Understanding the variability of these parameters across regions is crucial to answering our question, and to optimizing public health protocols. Furthermore, analyzing the trends in the SIR model can help us solve questions such as: "Which vaccine allocation strategy is most equitable? How many hospitalizations? and How long is the latent period?" (Oliveira Roster, 2024).

The SIR model divides a population into three compartments: Susceptible (S), Infectious (I), and Recovered (R). Susceptible represents individuals who can contract the disease. Infectives represent individuals who are actively transmitting the disease. Finally, recovered individuals have gained immunity or are no longer transmitting the disease. There is a transition between the compartments which are influenced by two key parameters: Beta (β), and Gamma (γ). β represents the infection rate, and the probability of disease transmission per contact, and γ is the recovery rate, and the proportion of infected individuals recovering per unit of time. The model can be described by the equations

$$\frac{dS}{dt} = - \frac{r \log(1-c)SI}{N} \quad (1)$$

$$\frac{dI}{dt} = - \frac{-r \log(1-c)SI}{N} - \gamma I \quad (2)$$

$$\frac{dR}{dt} = \gamma I \quad (3)$$

dX/dt means the rate of change of function X over time t . N is the total population, β is the transmission rate, γ is the recovery rate, and r is the rate of contact among individuals in the populations (Keeling et al., 2008) (Bank Group, 2021) (Gill, 2015).

The basic reproduction number (R_0) is the most significant part of the SIR model which is

$$R_0 = \beta/\gamma \quad (4)$$

This indicates the average number of secondary infections caused by one infected person in a fully susceptible population. $R_0 > 1$, the disease spreads; if $R_0 < 1$, the outbreak eventually dies out.

The COVID-19 pandemic has demonstrated the importance of understanding how infectious diseases spread and how interventions can mitigate the spread. Even with the widespread use of mathematical models such as the SIR, SEIS (Zhang et al., 2013), SEIR (Hou & Hoda Bidkhori, 2024), SIS (4.2: *The SIS Model*, 2022), and other mathematical models, significant gaps remain in our understanding of regional and national variations in the parameters. Understanding the gaps is essential to improving pandemic preparedness, tailoring responses to local contexts, and saving lives in the process.

In this study, we aim to clarify how the estimated SIR model parameters, specifically the infection rate (β) and recovery rate (γ), vary across different countries and regions. We want to identify the most important factors and interventions that help slow the spread of the virus. These factors include public health policies, healthcare infrastructure, vaccination rates, and social behaviour. Furthermore, we want to try to quantify the impact the specific interventions such as lockdowns, mask mandates, and vaccination campaigns have on infection and recovery rates. By analyzing their contribution we can have a more comprehensive understanding of our tools against the next virus.

This study will analyze the variation in the estimated SIR parameters (β and γ) across different countries and determine the impact of public health measures on the parameters. Our hypothesis, countries with larger populations would have higher β values due to increased opportunities for transmission and lower γ values, due to disparities in healthcare accessibility and response time. Additionally, we believed that quarantine measures would be most effective in reducing β , while vaccination programs would be necessary in improving γ values and overall controlling the outbreak.

In the Methods section, we explain the data sources, preprocessing steps, and the applications of the SIR model to estimate the parameters across six different countries. The Results section discusses the trends in infection, β , γ , and R_0 for each country, highlighting trends and possible correlations with health interventions. In the Discussion section, we use the findings, showing the role of population density, healthcare capacity, and policy measures in shaping trends, which identify limitations in the research. The Conclusion section summarises the research, illustrating the need for the SIR framework for future pandemic planning, future research directions, and the inclusion of socio-political variables and advanced modelling techniques.

Methods

The goal is to analyze the COVID-19 pandemic using the SIR model, focusing on the variation in the infection and recovery rates across countries and the impact of interventions. Our data processing can be broken down into 4 steps. First, we analyze the data source by using the COVID-19 dataset from Our World in Data (Mathieu et al., 2024).

Secondly, we cleaned up the data and prepared it using Python. We created a function that calculates the difference between consecutive cumulative case counts to get the weekly increase, due to the initial data having weekly frequency. We also broke the data down into distinct waves of the pandemic. We used a function to locate the local minima and maxima in the time series data. These points helped chop up the time series into individual waves so we could analyze them further.

Thirdly, we used the SIR model. We first calculated the SIR components. The daily new cases data was then used to calculate the S, I, and R components of the SIR model using a function. This function estimates S, I, and R based on the daily new cases and an assumed recovery period. We used another function to estimate the β and γ values, from the calculated S, I, and R values. This function uses S, I, and R changes over time to derive estimates for β and γ .

After that, we decided to select six countries to investigate. These five countries include Brazil, the United States, India, China, Russia, and Canada, due to their size, global impact, and proximity to the writers. We calculated γ and β for each week for each of the countries. We compared the estimated β and γ values between countries to determine their variation.

Results

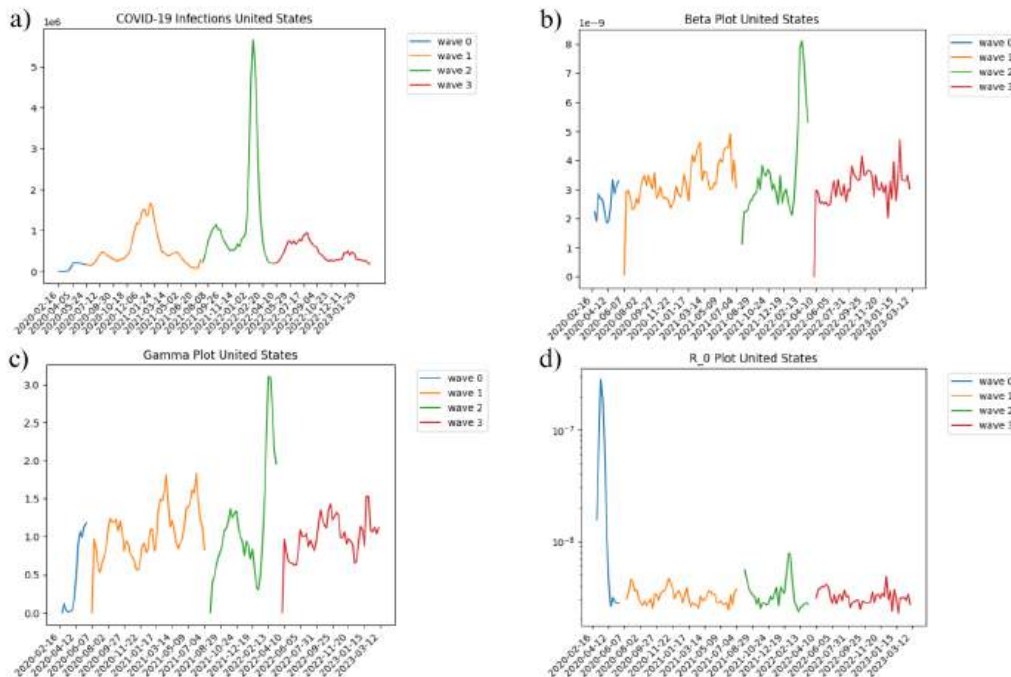


Figure 1: Trends in COVID-19 Infections and Epidemiological Parameters in the United States. Panels (a)-(d) display key epidemiological trends of COVID-19 in the United States across multiple waves. (a) Weekly reported infections, (b) estimated transmission rates, (c) estimated recovery rates, and (d) estimated basic reproduction number. Each wave is colour-coded: Wave0 (blue), Wave 1 (orange), Wave 2 (green), and Wave 3 (red).

After analyzing the SIR model of the United States of America, we can observe two major peaks in infections both in Wave 1 according to Figure 1a. In the US there were a total of 3 major waves. Wave 1 highlighted in orange had the biggest spike, which suggests a new variant of the virus or a lack of control measures. Wave 0 highlighted in blue and Wave 2 highlighted in green show much shorter peaks of infections. Wave 0 suggests the beginning of the virus, and Wave 2 shows the later stages of the now-controlled virus. γ represents the recovery rate. During Wave 0 of Figure 1b, we can see the spike of the γ value near the end of the wave, this is correlated to the stagnation of the cases at the beginning of Wave 1. The γ value varies between 0.5 and 2 throughout Wave 1. However, there is a sharp increase, nearing the end of Wave 1. The reason for this spike could be the rollout of vaccines which help limit the spread and lessen the symptoms of the disease. β values represent the infection rate. Wave 1 highlights the highest transmission rates, and the spike in transmission rates directly correlates with the spike in cases in Figure 1. Waves 0 and 2 in Figure 1b show lower and more stable β values and hover around $2-4 \times 10^{-9}$ range. R_0 (Basic reproduction number) is the average number of secondary infections caused by an infected individual. The reproduction number was highest at the beginning of Wave 0 since it was a new virus and there were not many measures. Over time, the R_0 value stabilizes for Waves 1 and 2 as we get a better understanding of the virus and due to interventions reduce the spread of the virus.

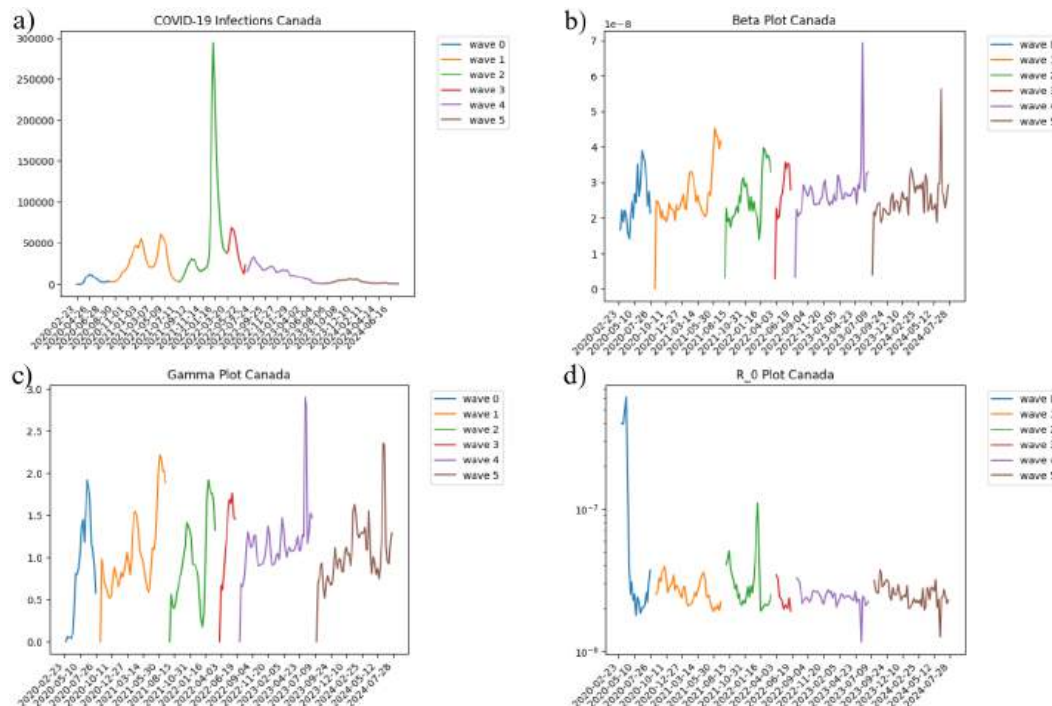


Figure 2: Trends in COVID-19 Infections and Epidemiological Parameters in Canada. Panels (a)-(d) display key epidemiological trends of COVID-19 in Canada across multiple waves. (a) Weekly reported infections, (b) estimated transmission rates, (c) estimated recovery rates, and (d) estimated basic reproduction number. Each wave is colour-coded: Wave 0 (blue), Wave 1 (orange), Wave 2 (green), Wave 3 (red), Wave 4 (purple), Wave 5 (brown).

Figure 2a shows 6 different waves of infections (Waves 0-5) from March 2020 to mid-2024. Wave 0 shows the earliest phase of the pandemic in spring-summer 2020, where infections begin at a relatively low rate, and then slowly rise. On March 21, 2020 (Wave 0), the Canadian Government set up restrictions on non-essential travel across the border and on April 15, 2020, enforced quarantine on anyone arriving in the country (Canada, 2020). Wave 1 in late 2020, experienced a sharper rise in infections relative to wave 0, which is reflected with increased community spread. Around Wave 1 COVID-19 vaccination began in Canada (Public Health Agency of Canada, 2021). "On November 26, 2021, the World Health Organization classified B.1.1.529 as a variant of concern called Omicron" (Canada, 2021). Due to the new variant Wave 2, which had the largest surge in the entire timeline (~300,000 cases at its maxima), this suggests a newer more transmissible strain or reduced compliance with public health measures. Wave 3, is more moderate than Wave 2 but still has a discernible bump in new infections. Wave 4 has a short spike when compared to Waves 2 and 3, however, it still is substantial. Wave 5 is from early 2022 to 2024, and shows a low plateau rather than any large peaks, this illustrates that overall infections had fallen off, likely due to roll out of vaccinations, accumulated immunity, or persistent public health measures.

Figure 2b is the estimated infection β for each wave. Although the β values are displayed in scientific notation (on the order of 10^{-8}), we will analyze the relative peaks and troughs over the years. β (on the order of $2-3 \times 10^{-8}$) starts at a moderate level, this shows the early stages of the virus, when there were fewer infections overall, and little to no immunity. As Canada adopted early interventions (border controls, initial social distancing measures), β remains in a moderate range without any significant peaks. During Wave 1, the β values seem to decrease initially, then trend upwards. This wave overlaps with late 2020, when stricter measures were relaxed in some provinces, and the virus might have spread more efficiently indoors during colder weather and decreases in sunlight and/or UV exposure (Rohde, 2021). Wave 2 shows Canada's largest infection spike in Figure 2a. The peak β value in Wave 2 is high, however, is not the highest of all waves. This underscores that while β was substantial (indicating easy transmission), the massive wave also owed itself to a relatively lower recovery rate or delayed recovery (see the γ discussion below). β peaks are relatively smaller than Wave 2. This could be due to partial herd immunity from the previous wave or renewed provincial restrictions. The P_{crit} (minimum proportion of the population required to achieve herd immunity) for Canada is 56.5%, this means that at least 56.5% of Canada's total population needs to become immune to COVID-19 (KWOK et al., 2020). In Wave 4 there is a significant spike in β . Even if the infections in Figure 2a do not reach Wave 2's magnitude, the β value in Wave 4 surpasses that of Wave 2. This is an indication of the virus being highly transmissible during this period. The spike in transmissibility could be due to the advent of the Omicron variant, but the shorter, sharper wave suggests that people were recovering faster or had natural immunity. During Wave 5, β values fluctuate but do not stay as high as in Wave 4, due to the combined effects of herd immunity, continued vaccinations, and milder variants.

Figure 2c shows how the γ (recovery rate) evolves over the 6 waves. During Wave 0 γ begins around 1.0-1.5 and climbs slightly. In early 2020, clinical protocols improved quickly, so even if hospitals were overused, the fraction of people recovering each day rose. During Wave 1 γ rises above 2.0 at times, suggesting faster recovery or shorter durations of the infection. This reflects better case management, higher recovery times (possibly due to herd immunity), increased use of therapies like corticosteroids (Monreal et al., 2020), or younger demographics being infected. The γ values dip closer to ~ 1.0 during Wave 2. This means a longer infectious period on average, which partly explains why Wave 2's infection curve peaked. One of many reasons for this could be that Hospitals could be overwhelmed with new cases, or more vulnerable populations were affected which increased the average recovery time. Wave 3 and 4 both exhibit rising γ with Wave 4 peaking around 2.5-3.0. This jump coincides with improved treatments and widespread vaccinations. Even if the virus was highly transmissible, the faster recovery rate in Wave 4 helped shorten the wave's overall duration, and infections. In Wave 5 γ remains above ~ 1.0 but is more variable. Many Canadians had either partial immunity from vaccines or a prior infection, so recovery was relatively swift.

Figure 2d is on a logarithmic scale. During the earliest part of the pandemic, there is a high spike in the R_0 . Although the absolute values appear small due to the plotting convention, the relative chances indicate that R_0 was largest when COVID-19 was new and no interventions were made. During waves 1 and 2, R_0 remains elevated with Wave 2's values spiking. This aligns with the large infection surge. Despite moderate β values in Wave 2, the notably lower γ would drive R_0 higher. During Wave 4, R_0 did not remain as high as Wave 2 for long due to the high β spike and the high γ spike coinciding. This indicates that although the virus was highly transmissible, faster recoveries or shorter infectious periods limited the wave's overall spread. In Wave 5, R_0 was relatively modest, with few fluctuations, it tends to stay on the lower side suggesting by 2022-2024, interventions such as vaccinations, mask mandates and partial immunity kept R_0 near or below the threshold that would cause a spike in infections.

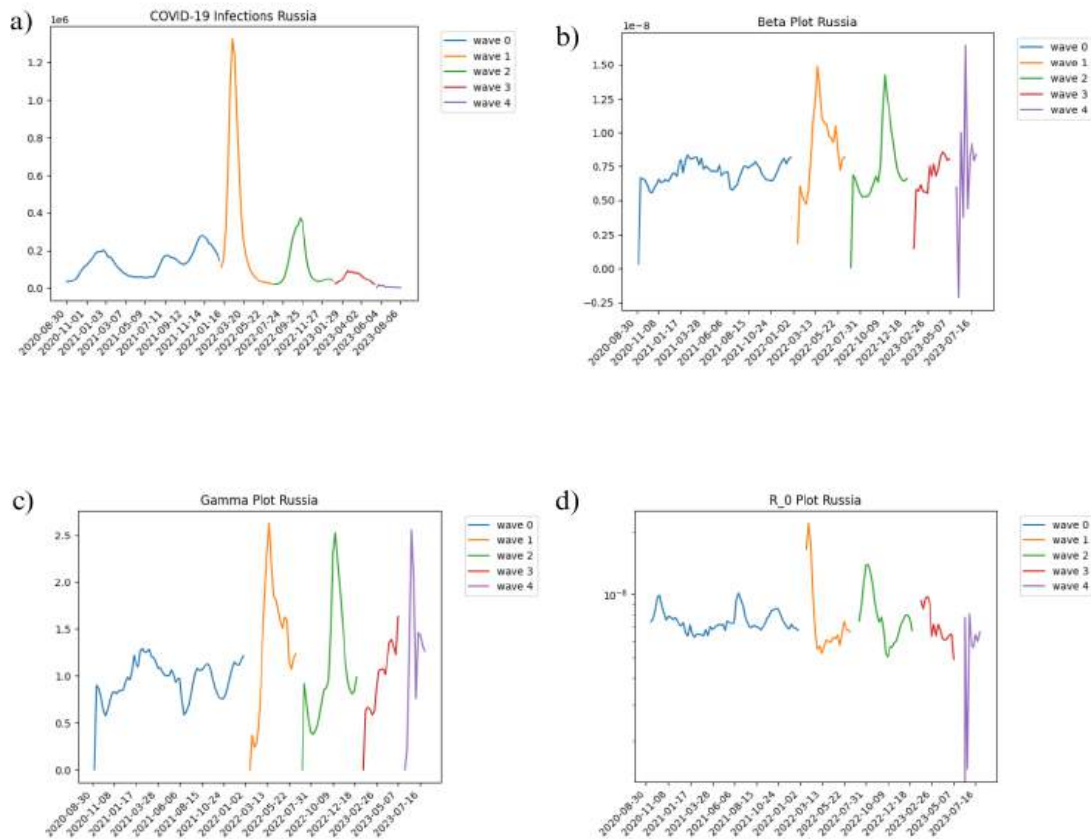


Figure 3: Trends in COVID-19 Infections and Epidemiological Parameters in Russia. Panels (a)-(d) display key epidemiological trends of COVID-19 in Russia across multiple waves. (a) Weekly reported infections, (b) estimated transmission rates, (c) estimated recovery rates, and (d) estimated basic reproduction number. Each wave is colour-coded: wave 0 (blue), Wave 1 (orange), Wave 2 (green), and Wave 3 (red).

In Figure 3a there seems to be a daily number of infections over time with distinct peaks representing different waves of the pandemic. Wave 1 has the highest peak, followed by smaller peaks for subsequent waves. In Figure 3b wave 0, β is stable, indicating moderate transmission. Toward the end of wave 0 however, there is a slight upward trend which predicts the second wave of the pandemic. A sharp spike in β values align with the infection peak in Figure 2a. This reflects heightened transmission during this wave, possibly because the disease has a more transmissible variant. Transmission rates in Waves 2 and 3 rise and fall in distinct patterns, though with smaller peaks than in Wave 1. This might indicate more control over the virus. In Wave 4 there is a rise in β values which could be because of waning immunity, emergence of new variants, or lessening of protocols. In Figure 3c γ remains relatively low and stable, reflecting consistent recovery, in wave 0. There is a sharp spike in γ which aligns with the infection peak in Figure 3a. This suggests a high number of recoveries during this period, possibly reflecting faster recovery in mild cases or increased healthcare response to the surge.

After the peak, however, γ falls back to the baseline levels, aligning with the end of Wave 1. In Figure 3d, R_0 fluctuates slightly above 1, indicating slow but persistent transmission during the initial phase. A dramatic spike in reproduction occurs in Wave 1, coinciding with the sharp infection rise in Figure 3a. This highlights the rapid disease spread during this wave. After the peak reproduction value drops below 1 reflecting effective control measures.

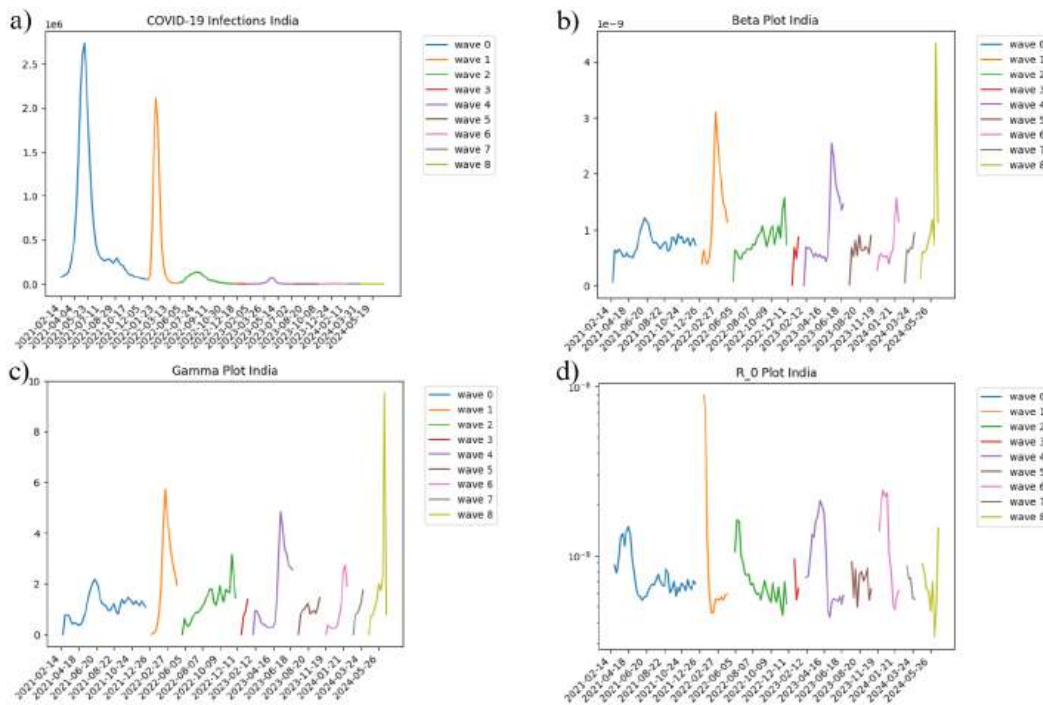


Figure 4: Trends in COVID-19 Infections and Epidemiological Parameters in India. Panels (a)-(d) display key epidemiological trends of COVID-19 in India across multiple waves. (a) Weekly reported infections, (b) estimated transmission rates, (c) estimated recovery rates, and (d) estimated basic reproduction number. Each wave is colour-coded: wave 0 (blue), Wave 1 (orange), Wave 2 (green), Wave 3 (red), Wave 4 (purple), Wave 5 (brown), Wave 6 (pink), Wave 7 (grey), Wave 8 (lime green).

India experienced nine distinct waves of COVID-19 infections from March 2020 to mid-2024. Wave 0 (March-June 2020) reflects the early stages of the virus, with infections rising due to strict nationwide lockdowns imposed on March 25 2020 (Embassy of India, Kyiv, Ukraine, 2023). However, limited accessibility to testing possibly underestimated the true case counts. At the end of wave 0 however, there was the largest spike in cases. The second largest spike occurred in Wave 1 (April-June 2021), peaking at over 1 million weekly cases. This aligns with the Delta variant's dominance (Nazario, 2022), which likely overwhelmed the healthcare system and caused widespread oxygen shortages (Welle (www.dw.com), 2021). Waves 3-8 had relatively smaller peaks when compared to Waves 0 and 1, and typically plateaued off.

In the earliest phase in Figure 4b, β hovers at a moderate level. This corresponds to India's strict nationwide lockdown from March 25, 2020 (Embassy of India, Kyiv, Ukraine, 2023). Lockdowns are expected to drive β values sharply downwards, but limited testing and errors in reporting during this early stage can cause β estimates to fluctuate, rather than plummet. Wave 1 shows a sharp spike in β values exceeding 2×10^{-9} . This aligns with the Delta variant's surge (Nazario, 2022). Delta's higher transmissibility likely contributed to the rapid rise in β values. Overwhelmed healthcare systems, especially the oxygen shortages (Welle (www.dw.com), 2021), forced the Indian government to re-impose some public health measures. These measures might have prevented the β values from climbing even higher. Following the Delta-driven peak, β values decline to more moderate levels (near 1×10^{-9}). Although the Delta variant was still rampant, partial immunity from prior infections and scaled-up vaccination efforts began to stagnate overall transmissions. The resulting β plateau suggests that, despite ongoing community spread, new waves were smaller than the massive Delta spike. In Waves 3-5, β fluctuates but rarely approaches the peaks of Wave 1. Periodic spikes (notably in the purple wave) suggest that either a sub-variant has occurred or the loosening of public health measures. These later waves show consistent spikes in β . Notably, Wave 8, has a sharp increase toward the end of the observed period (exceeding 4×10^{-9}). This jump could indicate either a more transmissible sub-variant, seasonal/behavioural changes, or relaxation of public health measures. By this time large-scale lockdowns were no longer common, so any spike in β would have been managed by the ongoing vaccination effort.

During Wave 0 in Figure 4c γ remains around 1-2, signifying that infected individuals took roughly a week to a week to recover on average (i.e., $1/\gamma$ in days). Early in the pandemic, limited treatments and testing constraints could skew these estimates. Furthermore, overworked hospitals might have slowed recoveries with severe cases, and the milder asymptomatic cases were often unreported, which might've elevated γ . Wave 1's γ spikes (surpassing 4 or 5 at certain points) correspond to a shorter infectious period. There are several factors to explain the high drive of γ during this wave. Firstly, selective testing of more symptomatic individuals, who might also isolate sooner. Additionally, improved medical protocols for moderate cases. During Wave 2, γ falls back toward 2, suggesting a longer infectious period relative to Wave 1. This could be due to overwhelmed hospitals. It is important to note that by late 2021, India was still coping with Delta's aftermath, and while vaccines were rolling out, they may not have been distributed evenly. In Waves 3-5 γ oscillates between ~ 1 and ~ 3 . Surges in γ often coincide with improved treatment availability or a variant that was milder due to herd immunity. γ dips could be represented by stressed-out healthcare systems. Later waves have a general uptrend in γ , with Wave 8 showing a significant jump above 8 or 9. Such a high γ implies very short average infectious durations. There are many factors which contribute to this general uptrend in γ values. Evolving medical treatments, widespread vaccination, and booster doses likely led to the increase in γ . Another possibility is that these late-wave cases were very mild or asymptomatic which further reduces the time to recover in official data.

In Wave 0 Figure 4d, R_0 fluctuates in the mid-range of 10^{-9} on the graph. This means that there is a moderate level of transmissibility relative to the recovery rate. Although India was under lockdown, a lack of following the rules and testing constraints allowed the virus to spread at a moderately slow rate. In Wave 1, there was a very large spike in R_0 visible, this means that β was high and γ was lower but still elevated. This spike correlates with the Delta variant which was both highly transmissible (raising β) and capable of quickly infecting susceptible hosts who, once infected, might recover or be removed from circulation. This caused the explosive outbreak during Wave 1. In Wave 2, R_0 settles at a lower value. After Wave 2, each wave has its own distinct bump (often when β rises or γ dips). For example, Wave 4 has a mild spike in R_0 , due to β climbing faster than γ . Waves 6-8, R_0 have moderate surges and dips, but none reach the same peak as Wave 1. Even though Wave 8's β surges, the corresponding γ is also extremely high, keeping the R_0 in check.

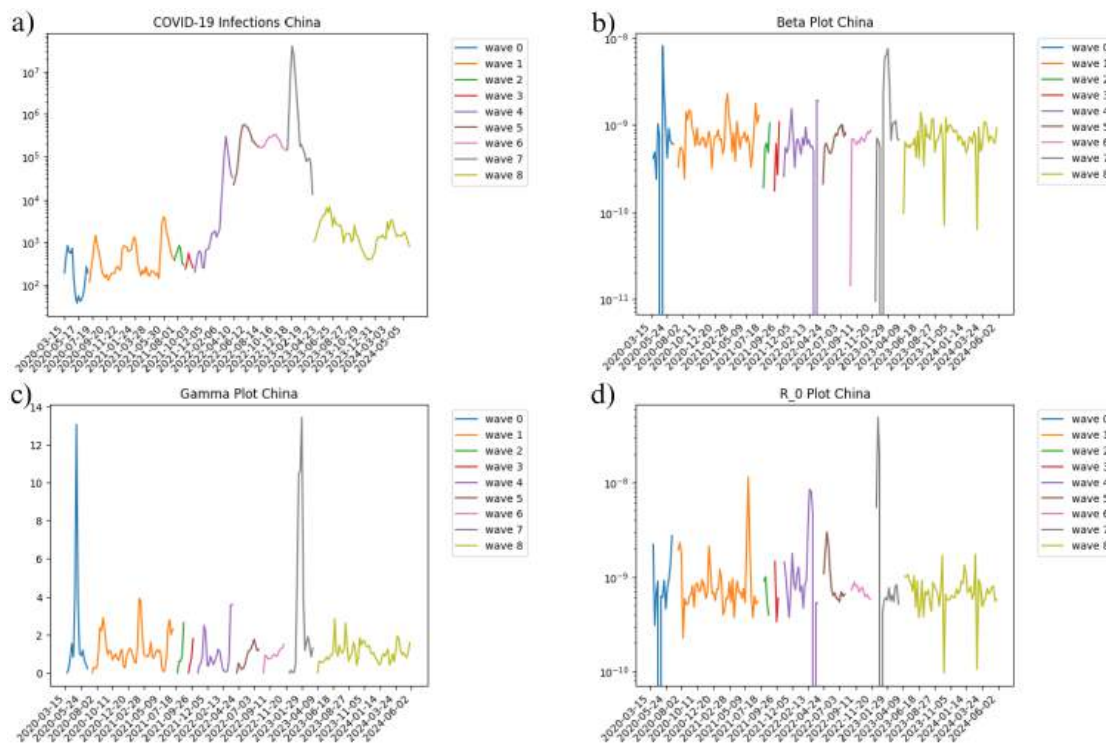


Figure 5: Trends in COVID-19 Infections and Epidemiological Parameters in China. Panels (a)-(d) display key epidemiological trends of COVID-19 in China across multiple waves. (a) Weekly reported infections, (b) estimated transmission rates, (c) estimated recovery rates, and (d) estimated basic reproduction number. Each wave is colour-coded: wave 0 (blue), Wave 1 (orange), Wave 2 (green), Wave 3 (red), Wave 4 (purple), Wave 5 (brown), Wave 6 (pink), Wave 7 (grey), Wave 8 (lime green).

China's COVID-19 response had strict lockdowns, mass testing, quarantine policies, and until late 2022 a "zero-COVID" strategy. Below are some well-known public health measures that likely impacted β , γ , and R_0 values. Wuhan Lockdown (January 23, 2020) was the first major lockdown imposed in Wuhan and nearby cities in the Hubei Province (BBC, 2020). Nationwide controls (Spring-Summer 2020) had a rapid expansion of lockdowns, mandatory mask-wearing, and social distancing. China frequently conducted city-wide or district-wide testing, quarantined contacts, and closed high-risk venues through 2020-2022 (Zhou et al., 2021), (Reuters Staff, 2020). In 2022 some regions in China experimented with lighter restrictions while others locked down due to Omicron sub-variants (Yeung, 2022). In early 2023, large-scale restrictions were lifted nationwide, leading to a surge in cases (Pei, 2023).

In Figure 5a, we can observe the number of cases over time. The infection curve for China is plotted on a logarithmic scale, ranging roughly from 10^1 to 10^7 . It has eight distinct colour-coded waves. Wave 0, is relatively low, likely reflecting the strict Wuhan lockdown and national travel restrictions in the early stages of the virus. Waves 1-3 show a stepwise increase in reported infections, but still below the peaks seen in many Western countries at the time. This indicates the effectiveness of the strict policies in containing the virus. Wave 6, had the largest peak, where there was a sudden surge in the grey wave suggesting that after the abrupt end of the zero-COVID in January 2023, infections spiked sharply after. Wave 7 and 8, appear at lower magnitudes than Wave 6, implying partial immunity or less severe variants.

In Figure 5b, there was an initial spike in β values in wave 0, which is likely due to early outbreaks in Wuhan before strict measures took full effect. During Waves 1 to 5, β values remain relatively stable reflecting China's strict control measures such as lockdowns, mass testing, and travel restrictions. In Wave 6, the transmission rate fluctuates significantly, which likely corresponds with the lifting of zero-COVID policies in early 2023, leading to a sharp rise in cases. Wave 8, shows more stabilized β values with some fluctuations, possibly indicating smaller outbreaks or new sub-variants.

In Figure 5c, γ is relatively low in early waves but spikes briefly in wave 0, possibly due to early hospitalization and improved containment of the diseases. Throughout waves 1-5, γ values remain relatively low and variable, likely due to prolonged hospital stays and strain on healthcare systems in certain regions. Wave 6 sees a sudden increase in recovery rates, possibly due to a milder Omicron variant, improved outpatient treatments, or reduced hospitalization criteria. Wave 8, shows more stable but moderate values, suggesting the endemic phase of the COVID-19 with controlled healthcare responses.

In Figure 5d, similar to β , wave 0 shows extreme fluctuations, representing early uncertainties and rapid response measures in R_0 . Wave 1-5 shows relatively moderate values, aligning with strict control policies keeping the virus in check. A sharp increase in Wave 6 reflects the abrupt policy shift in early 2023, leading to widespread transmission. Wave 8, stabilizes at lower values, which likely indicates immunity buildup, vaccinations, and less virulent strains.

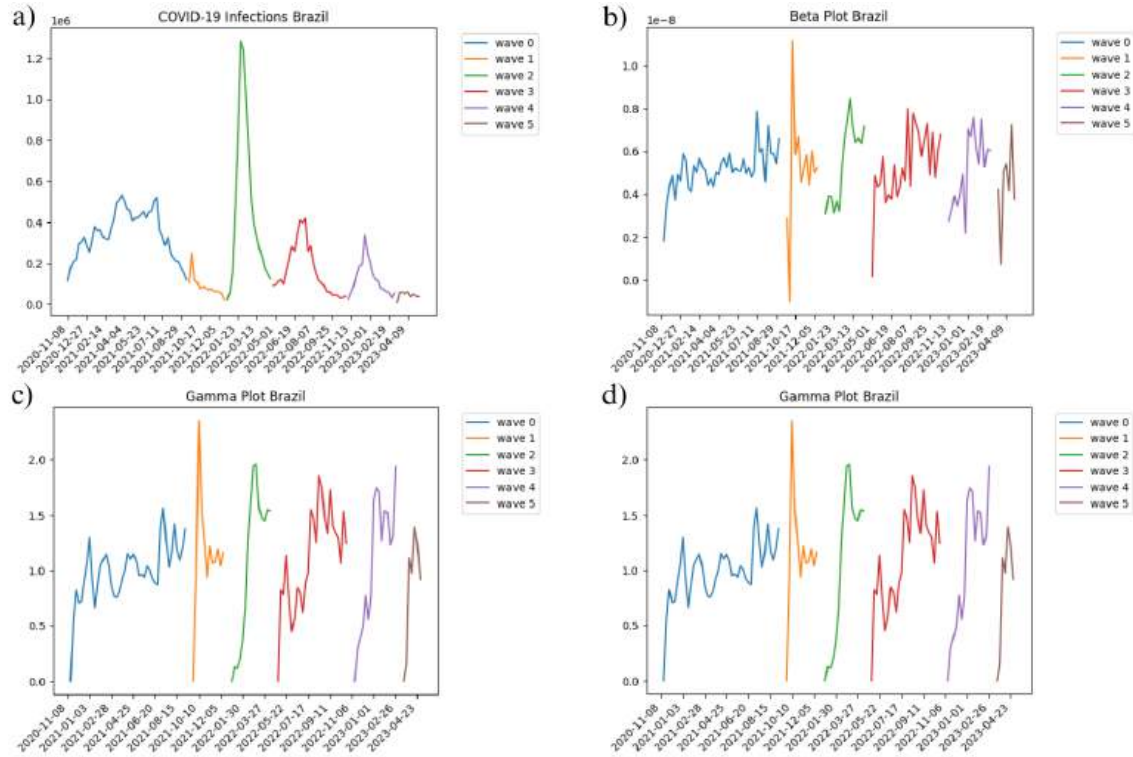


Figure 6: Trends in COVID-19 Infections and Epidemiological Parameters in Brazil. Panels (a)-(d) display key epidemiological trends of COVID-19 in China across multiple waves. (a) Weekly reported infections, (b) estimated transmission rates, (c) estimated recovery rates, and (d) estimated basic reproduction number. Each wave is colour-coded: wave 0 (blue), Wave 1 (orange), Wave 2 (green), Wave 3 (red), Wave 4 (purple), Wave 5 (brown).

In Figure 6a Wave 0, started in late 2020 and continued into early 2021. This period coincided with Brazil's early non-pharmaceutical interventions (NPIs), including state-level lockdowns and mask mandates implemented in mid-2020 (Bernardeau-Serra et al., 2021) (Library of Congress, 2020). However, lack of coordination in the response by the federal government contributed to the high infection rates. Wave 1, was short and by this time Brazil had launched its vaccination campaign prioritizing healthcare workers and vulnerable populations (Bernardeau-Serra et al., 2021). The introduction of variants like Gamma may have driven infections. Wave 2, had a massive spike, which likely correlates with the Omicron variant, which spread rapidly worldwide. Brazil had increased its vaccination coverage by this time, however, due to Omicron's high transmissibility the infection rates continued to rise. Wave 3, had a smaller wave, having a shorter spike than waves 0 and 2. During this wave, Brazil declared the end of most of its COVID-19 non-pharmaceutical interventions in April 2022 (Library of Congress, 2022). Waves 4 and 5, had shorter waves, gradually plateauing. The plateau in infection rates could be due to high vaccination rates and natural immunity.

In Figure 6b Wave 0, β values were moderate to high as the virus spread in a new susceptible population. Wave 1, β values were elevated due to the Gamma variant and the partial compliance with NPIs. It is highly unlikely that the rollout of the vaccinations had any effect on the β values as it was in the early stages of the vaccination rollout process. During Wave 2, β has

a notable increase, this could possibly correlate with Omicron's higher transmissibility, driving a steep rise in cases (Katella, 2023). During Wave 3, β still fluctuates but is generally lower than in early 2022, this shows the spread of the virus has been slowing down, due to possible accumulating of immunity. Wave 4 and 5, have generally lower, and less frequent spikes, indicating the late stages of the virus.

In Figure 6c and Figure 6d, there are similar trends to the β plot suggesting that recovery rates were influenced by the availability of medical resources. Higher γ values indicate improved treatments and hospital responses. Sustained γ values in the later waves illustrate the growth in medical treatments for the virus.

Discussion

Our results indicated that there is a variation in β and γ values across different countries, which aligns with our hypothesis. Countries such as Brazil and the United States experienced large spikes and had greater β values during the peaks. This suggests that densely populated regions or those with lower compliance with public health measures had higher transmission rates. Countries with more rapid healthcare infrastructure such as China and Canada, had lower β values and higher γ values, illustrated a faster R_0 .

The effectiveness of public health measures in controlling transmission was evident in our analysis. Lockdowns and social distancing rules correlated with the decline in β values across multiple regions. For example, in Brazil, the β value spiked before the early 2022 wave, which coincided with a new mutation. The γ values increased when the new vaccine was rolled out. A similar trend was seen in the United States, where the rollout of vaccines corresponded with a gradual decrease in β values and the stabilization of γ values. These findings showcase the importance of timely and widespread vaccination campaigns and effective isolation in mitigating the spread of infectious diseases.

One of the most important observations made in our study was the relationship between the basic reproduction number and public health interventions. Initially, R_0 values were highest in wave 0 across all studied countries due to the lack of herd immunity and minimal knowledge of how they contain the virus. With the implementation of control measures such as isolation and mask mandates, the R_0 value stabilized, particularly in countries with consistent health protocols.

Our study has several limitations. Firstly, there are many confounding factors which affect the spread of diseases. Future research needs to focus on enhancing the SIR model and making it more accurate, investigating the role of socioeconomic and political factors, exploring machine learning-based epidemiological models, and analyzing more diseases to find a greater trend.

The SIR model has some limitations as it assumes too many variables. SIR models can simulate short-term patterns, but their assumptions (like closed populations and permanent immunity) are not very accurate due to the world's complexities, such as behavioural changes, infections, and external influences. To better predict these outcomes, a more advanced model is

required to analyze the data, as the SIR Model is too simple. This study shows some of the inefficiency of SIR models in forecasting the COVID-19 epidemic (Moein et al., 2021). While the model can provide initial insights, it often oversimplifies complex dynamics such as asymptomatic transmission, heterogeneous contact structures, and vaccination effects.

Conclusion

This paper analyzed the application of the Susceptible-Infectious-Recovered (SIR) model to COVID-19 transmission patterns across different countries. By estimating key parameters such as the infection rate (β) and recovery rate (γ), we identified variations in disease spread and intervention effectiveness. Our findings suggest that factors like population density, healthcare capacity, and public health measures significantly impact these parameters, with interventions such as lockdowns and vaccination programs playing a crucial role in controlling outbreaks.

Several key questions emerged from our study. Can more complex models, such as SEIR or agent-based simulations, provide a more accurate representation of real-world disease dynamics? How do socio-political factors influence β and γ beyond traditional epidemiological variables? Future research could explore machine learning techniques to refine parameter estimation or investigate the long-term effects of interventions on viral evolution. Additionally, examining how different vaccination strategies impact γ over time could enhance pandemic preparedness.

By addressing these gaps, researchers and policymakers can develop more tailored, data-driven responses to future infectious disease outbreaks.

Acknowledgements

I would like to thank Plinio Zanini for their consistent feedback, and support throughout this research paper. I would also like to thank my family for their continuous support. Finally, I would like to thank the resources and tools which I used to research, including data from COVID-19 OurWorldData and the World Health Organization.

Data availability

Data and code for this study can be found at

<https://github.com/NipunS08/Covid-Spread-SIR-Modeling>.

Works Cited

- 4.2: *The SIS Model*. (2022, January 5). Mathematics LibreTexts.
[https://math.libretexts.org/Bookshelves/Applied_Mathematics/Mathematical_Biology_\(Chasnov\)/04:_Infectious_Disease_Modeling/4.02:_The_SIS_Model](https://math.libretexts.org/Bookshelves/Applied_Mathematics/Mathematical_Biology_(Chasnov)/04:_Infectious_Disease_Modeling/4.02:_The_SIS_Model)
- Ba, Z., Li, Y., Ma, J., Qin, Y., Tian, J., Meng, Y., Yi, J., Zhang, Y., & Chen, F. (2023). Reflections on the dynamic zero-COVID policy in China. *Preventive Medicine Reports*, 36, 102466–102466. <https://doi.org/10.1016/j.pmedr.2023.102466>
- Bank Group, W. (2021, January 1). *An introduction to deterministic infectious disease models*. World Bank Group. <https://coilink.org/20.500.12592/s83912>
- Bernardeau-Serra, L., Nguyen-Huynh, A., Sponagel, L., Sernizon Guimarães, N., Teixeira de Aguiar, R. A., & Soriano Marcolino, M. (2021). The COVID-19 Vaccination Strategy in Brazil—A Case Study. *Epidemiologia*, 2(3), 338–359.
<https://doi.org/10.3390/epidemiologia2030026>
- Calfas, J. (2020, October 15). *Why Covid-19 Is Spreading Again: Fatigue, Colder Weather, Eased Restrictions*. WSJ; The Wall Street Journal.
https://www.wsj.com/articles/why-covid-19-is-spreading-again-fatigue-colder-weather-eased-restrictions-11602759601?reflink=desktopwebshare_permalink
- Canada, P. H. A. of. (2021, November 30). *Government of Canada introduces additional measures to address COVID-19 Omicron variant of concern*. Wwww.canada.ca.
<https://www.canada.ca/en/public-health/news/2021/11/government-of-canada-introduces-additional-measures-to-address-covid-19-omicron-variant-of-concern.html>
- Canada, P. S. (2020, November 27). *Mandatory Quarantine*. Wwww.publicsafety.gc.ca.
<https://www.publicsafety.gc.ca/cnt/trnsprnc/brfng-mtrls/prlmntry-bndrs/20201201/010/index-en.aspx>
- COVID-19 Epidemiological Update*. (n.d.).
https://cdn.who.int/media/docs/default-source/documents/emergencies/20241224_covid-19_epi_update_special-edition.pdf?sfvrsn=b0a6ddaf_1&download=true
- Embassy of India, Kyiv, Ukraine*. (2023). Eoiukraine.gov.in.
<https://www.eoiukraine.gov.in/lockdown.php>
- Gill, N. (2015). *DETERMINISTIC AND STOCHASTIC MODELS OF INFECTIOUS DISEASE: CIRCULAR MIGRATIONS AND HIV TRANSMISSION DYNAMICS*.
<https://math.uchicago.edu/~may/REU2015/REUPapers/Gill.pdf>
- Government Of Canada. (2024). *Interactive Data Visualization of COVID-19 in Canada - Public Health Infobase | Public Health Agency of Canada*. Health-Infobase.canada.ca.
<https://health-infobase.canada.ca/covid-19/>
- Harko, T., Lobo, F. S. N., & Mak, M. K. (2014). Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation*, 236, 184–194.
<https://doi.org/10.1016/j.amc.2014.03.030>
- Hou, Y., & Hoda Bidkhori. (2024). Multi-feature SEIR model for epidemic analysis and vaccine prioritization. *PloS One*, 19(3), e0298932–e0298932.
<https://doi.org/10.1371/journal.pone.0298932>

- Katella, K. (2023, September 1). *Omicron, delta, alpha, and more: What to know about the coronavirus variants*. Yale Medicine; Yale Medicine.
<https://www.yalemedicine.org/news/covid-19-variants-of-concern-omicron>
- Keeling, M., Rohani, P., & Pourbohloul, B. (2008). Modeling Infectious Diseases in Humans and Animals: Modeling Infectious Diseases in Humans and Animals. *Clinical Infectious Diseases*, 47(6), 864–865. <https://doi.org/10.1086/591197>
- KWOK, K. O., LAI, F., WEI, W. I., WONG, S. Y. S., & TANG, J. (2020). Herd immunity – estimating the level required to halt the COVID-19 epidemics in affected countries. *Journal of Infection*, 80(6). <https://doi.org/10.1016/j.jinf.2020.03.027>
- Library of Congress, T. (2020, July 7). *Brazil: New law makes use of masks mandatory in some locations*. Library of Congress, Washington, D.C. 20540 USA; The Library of Congress.
<https://www.loc.gov/item/global-legal-monitor/2020-07-10/brazil-new-law-makes-use-of-masks-mandatory-in-some-locations/>
- Library of Congress, T. (2022, May 2). *Brazil: Ministry of health ends public health emergency*. The Library of Congress; The Library of Congress.
<https://www.loc.gov/item/global-legal-monitor/2022-05-02/brazil-ministry-of-health-ends-public-health-emergency/>
- Lockdowns rise as China tries to control virus. (2020, January 23). *BBC News*.
<https://www.bbc.com/news/world-asia-china-51217455>
- Mathieu, E., Ritchie, H., Rod s-Guirao, L., Appel, C., Gavrilov, D., Giattino, C., Hasell, J., Macdonald, B., Saloni Dattani, Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2024, November 17). *COVID-19 Pandemic*. Our World in Data.
<https://ourworldindata.org/coronavirus#article-citation>
- Melikechi, O., Young, A. S., Tang, T., Bowman, T. S., Dunson, D., & Johndrow, J. E. (2022). Limits of epidemic prediction using SIR models. *Journal of Mathematical Biology*, 85(4). <https://doi.org/10.1007/s00285-022-01804-5>
- Moein, S., Nickaeen, N., Roointan, A., Borhani, N., Heidary, Z., Javanmard, S. H., Ghaisari, J., & Gheisari, Y. (2021). Inefficiency of SIR models in forecasting COVID-19 epidemic: a case study of isfahan. *Scientific Reports*, 11. <https://doi.org/10.1038/s41598-021-84055-6>
- Monreal, E., Sainz, S., Natera-Villalba, E.,  lvaro Beltr n-Corbellini, Rodr guez-Jorge, F., Fern ndez-Velasco, J. I., Walo-Delgado, P., Muriel, A., Zamora, J., Alonso-Canovas, A., Jes s Fort n, Manzano, L., Montero-Errasqu n, B., Costa-Frossard, L., Masjuan, J., Villar, L. M., L. M  iz-Carro, S nchez-Garc a, E. M., Hidalgo, F., & Dom nguez, A. R. (2020). High versus standard doses of corticosteroids in severe COVID-19: a retrospective cohort study. *European Journal of Clinical Microbiology & Infectious Diseases*, 40(4), 761–769. <https://doi.org/10.1007/s10096-020-04078-1>
- Nazario, B. (2022, December 26). *Coronavirus (COVID-19) Overview*. WebMD.
<https://www.webmd.com/covid/coronavirus>
- Pei, M. (2023, March 2). *The Sudden End of Zero-Covid: An Investigation*. China Leadership.
<https://www.prclleader.org/post/the-sudden-end-of-zero-covid-an-investigation>

- Prodanov, D. (2022). Analytical solutions and parameter estimation of the SIR epidemic model. *Mathematical Analysis of Infectious Diseases*, 163–189.
<https://doi.org/10.1016/B978-0-32-390504-6.00015-2>
- Public Health Agency of Canada. (2019). *National Advisory Committee on Immunization - Canada.ca*. Canada.ca.
<https://www.canada.ca/en/public-health/services/immunization/national-advisory-committee-on-immunization-naci.html>
- Public Health Agency of Canada. (2021, January 15). *Demographics: COVID-19 vaccination coverage in Canada - Canada.ca*. Government of Canada.
<https://health-infobase.canada.ca/covid-19/vaccination-coverage/>
- Reuters Staff. (2020, October 12). China's Qingdao orders city-wide testing after new COVID-19 infections. *Reuters*.
<https://www.reuters.com/article/business/healthcare-pharmaceuticals/chinas-qingdao-orders-city-wide-testing-after-new-covid-19-infections-idUSKBN26X01S/>
- Rohde, R. (2021, January 8). *Colder weather and increasing coronavirus (COVID-19) spread - Berkeley Earth*. Berkeley Earth.
<https://berkeleyearth.org/colder-weather-and-increasing-coronavirus-covid-19-spread/>
- Statistics Canada. (2022, November 23). *Impacts of COVID-19 - Canada at a Glance, 2022*.
Www150.Statcan.gc.ca. <https://www150.statcan.gc.ca/n1/pub/12-581-x/2022001/sec14-eng.htm>
- Talic, S., Shah, S., Wild, H., Gasevic, D., Maharaj, A., Ademi, Z., Li, X., Xu, W., Mesa-Eguiaagaray, I., Rostron, J., Theodoratou, E., Zhang, X., Motee, A., Liew, D., & Ilic, D. (2021). Effectiveness of Public Health Measures in Reducing the Incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: Systematic Review and meta-analysis. *BMJ*, 375(8315). <https://doi.org/10.1136/bmj-2021-068302>
- The Lancet Regional Health – Western Pacific. (2023). The end of zero-COVID-19 policy is not the end of COVID-19 for china. *The Lancet Regional Health - Western Pacific*, 30, 100702. <https://doi.org/10.1016/j.lanwpc.2023.100702>
- U.S. Department of Health and Human Services. (2020, December 12). *COVID-19 Vaccines*. HHS.gov. <https://www.hhs.gov/coronavirus/covid-19-vaccines/index.html>
- Welle (www.dw.com), D. (2021, April 5). *COVID: Why is India facing an oxygen shortage? | DW | 04.05.2021*. DW.COM. <https://www.dw.com/en/india-covid-oxygen-shortage/a-57425951>
- Yeung, J. (2022, December 7). *China scraps some of its most controversial Covid rules, in significant step toward reopening*. CNN.
<https://www.cnn.com/2022/12/07/china/china-covid-restrictions-explainer-intl-hnk/index.html>
- Zhang, H., Yingqi, L., & Xu, W. (2013). Global stability of an SEIS epidemic model with general saturation incidence. *ISRN Applied Mathematics*, 2013, 1–11.
<https://doi.org/10.1155/2013/710643>
- Zhou, L., Seesagur, H. N., Akhtar, N., Boolakee, J., & Pratt, C. B. (2021). A citywide “virus testing”: Chinese government’s response to preventing and controlling the second outbreak of sars-cov-2. *Frontiers in Public Health*, 9.
<https://doi.org/10.3389/fpubh.2021.601592>

Rethinking Education for Foster Children: Addressing Struggles Through Comprehensive Educational Strategies By Jolie Zhou

Achievement gap – the discrepancy in academic performance between foster youth and their non-foster peers

Aging out – a foster youth reaching the maximum age of foster care and subsequently exiting the system

Chronic absenteeism – the number of students absent 10% or more school days in the year

Disproportionate discipline – a group of students having a higher or lower rate of discipline, often represented through suspensions, expulsions, and other disciplinary actions

Extended foster care – programs that allow youth to remain in the system beyond the age of eighteen

Foster youth or foster child – a minor that has been placed in the care of a state-certified caregiver

Free and reduced-price lunch (FRPL) – a federal program that supports low-income students with meals

Independent living program (ILP) – a program intending to assist a foster youth's transition to adulthood

Individualized education program (IEP) – a customized and formal educational plan for a child with a disability that includes special education instruction, supports, and services

Kinship care – a foster care placement with a relative foster family

School instability – an event that disrupts a student's ability to learn such as a school transfer

School mobility – a voluntary or involuntary change in school placement outside of the typical grade transition

School of origin – the school in which the child was last enrolled or attended when permanently housed, or any school that a child feels connected to and has attended within the last fifteen months

Stability rates – the total count of cumulatively enrolled students with an enrollment start date on or after July 1st and on or before June 30th

Transition-age youth (TAY) – foster youth between the ages of sixteen and twenty-four that will soon leave the system

Trauma-informed teaching – strategies that support students that have experienced trauma

Abstract

Due to the nature of the foster care system, foster youth inevitably encounter vastly different experiences than non-foster youth. Notably, frequent changes in home and school placements significantly influence a foster child's ability to learn. Multiple sources have observed a discrepancy in academic performance, disciplinary behavior, and various social skills in foster youth. Detrimentially, these factors only accumulate over time, widening the performance gap in the long run. Many of the youth not only struggle to find meaning in education, but are also unsure about their future plans. This insists on the need for effective resources and support systems that address the unique needs of foster youth. For instance, educators that work directly with foster children should understand the system and be trauma-informed to best support their needs. However, a myriad of current approaches in the field are lacking in evidence-based strategies and research. To that end, this paper hopes to trailblaze the educational sector for foster youth, analyzing potentially beneficial projects and implementing initiatives that prove to be impactful. Through interviews with credible individuals, such as those from the district, County Office of Education, and nonprofit organizations, the common challenges of foster youth in education are learned through anecdotal tellings.

Introduction

Foster care poses an unorthodox developmental environment for children. According to the 29th Adoption and Foster Care Analysis and Reporting System (AFCARS) administered by the United States Department of Health and Human Services, 606,031 foster children were served by the system during the fiscal year of 2022 (AFCARS, 2022). The United Way of the National Capital Area explains the foster care system: it functions as a temporary arrangement for children whose biological parents are unable to meet their basic needs. Oftentimes, these parents go through emergency situations like incarceration, eviction, mental health issues, or addiction. Foster children can be adopted by their foster homes or move from home to home. Congregate care such as group homes are also an option for foster children in similar situations, allowing children with similar experiences to live together under supervision. Local, state, and federal agencies work together within this complex system to provide safe environments for these children. The Child and Family Services Agency (CFSA) decides whether or not to remove children from their home through evidence-based strategies. However, the CFSA prioritizes placing the children in kinship homes with an ultimate goal of reuniting the children with their biological parents. In creating a safe and healthy space for foster children to grow, foster parents have an age, citizenship, and minimum number of training hours requirement. Criminal and background checks are also completed to ensure the safety of the residence. The state is expected to provide varied amounts of financial assistance for the basic costs of raising a child, although iFoster¹, a nonprofit that provides resources for individuals in the foster care system, notes that these numbers are less than half of the average amount a family spends for their children.

According to Kools et al. (1997), the foster child experience has become an increasingly permanent living situation. While policies like the Adoption and Safe Families Act (ASFA) have attempted to decrease the waiting time and increase adoption options, the AFCARS national dataset shows that the number of children that exited foster care during the fiscal year decreased from 237,000 in 2013 to 201,000 in 2022. On average, the number of children waiting for adoption has increased from 102,000 to 109,000. If children are not or have not yet been adopted, they return to temporary care, long-term foster care, independent living, or permanent foster placements. Chipungu et al. (2004) [1] corroborates that foster children have unique developmental needs and are aware of their differences as soon as age two. The participants studied in the paper began to realize the social meaning embedded in these differences, which ensued as negative self-images and racial labels. Similarly, Denuwelaere et al. (2007) [2] finds that conflict level with foster parents is strongly correlated with depressive symptoms and self-esteem in foster children. Grounded in the belief that foster children will only reside temporarily, foster parents struggle with the dilemma of support. Denuwelaere writes that foster parents desire to foster a supportive environment, but the notion that more often than not, the children will leave their home, clouds their judgement. In parallel, Coleman et al. (1997) supports this thought, finding that adoptive families have relationships with more distance and conflict than birth families.

¹ <https://www.ifoster.org/blogs/how-does-the-foster-care-system-work/>

Education in foster care presents unique challenges for children's academic development. For instance, previous research shows that foster children have lower educational attainment, lower graduation rates, lower college participation, and higher participation in special education. According to Somers et al. [3], foster children are twice as likely to have an out-of-school suspension, have a higher rate of referrals and grade retention, and score lower on state assessments. Additionally, the National Working Group on Foster Care and Education found that only 50% of foster youth graduate high school by age 18, compared to 84% of their peers. The instability inherent in the foster care system often disrupts children's education, making it difficult to maintain academic progress or develop a strong educational identity. Foster children may struggle with motivation due to numerous commitments interfering with their studies, such as court appearances, therapy sessions, and family visitations. These obligations, along with frequent placement changes, can disrupt school attendance and study routines.

Educational Experience or Outcome		Findings
		National/Multistate
School Stability	% of youth in foster care who change schools when first entering care	31%-75% ¹
	% of 17-18 year-olds who experienced 5 or more school changes	25%-34.2% ²
School Engagement	% chronically absent from school	About twice the rate of non-foster students ³
	% of 17-to-18 year-old youth in foster care having out-of-school suspensions	12%-23% (compared to 5-7% of all students) ⁴
	% of 17-to-18 year-old youth in foster care being expelled	3-4 times that of non-foster students ⁵
Reading Attainment	Reading level of 17-to-18 year-old youth in foster care	29% -33% (meet state standards) ⁶
Special Education	% of youth in foster care receiving special education services	30%-50% ⁷ (compared to 14% for all students ⁸)
High School Graduation	% of youth in foster care who complete high school by age 18 (via a diploma or GED)	64% of foster youth compared to 87.3% for non-foster youth ⁹
Postsecondary Education	% of 17-to-18-year-old youth in foster care who want to go to college	70% ¹⁰ - 84% ¹¹
	% of youth in foster care who graduated high school who enrolled in postsecondary education at some level	13%-38% ¹²
	% of foster care alumni who attain a bachelor's degree	2% ¹³ -10.8% ¹⁴

Figure 1. Statistics about foster children in education revealing the issues and discrepancies between non-foster children [4]

The American Bar Association's 2022 foster care education fact sheet [5] pinpointed commonalities for foster children in education. Frequent school changes result in lower standardized testing scores and a greater risk of dropping out of school. The less placement changes a student experiences, the more likely they are to graduate. In fact, one study cited in the fact sheet found that "youth who had even one fewer change in living arrangement per year were almost twice as likely to graduate from high school before leaving foster care." These school changes not only affect a child's academic performance, but also renders it difficult to form and maintain relationships. Without sustained relationships, the resilience, well-being, and health of foster children were found to be affected. Oftentimes, foster children experience enrollment delays, the repetition of courses, lack of support with their special education needs, and

inappropriate course selection. While absenteeism is generally high for foster children, age, pre-foster care experiences and events during foster care were labeled as important factors in determining these rates. For the most part, suspension and expulsion rates were also higher for foster children than their peers. Figure 2 shows the tendency for foster children to experience higher chronic absenteeism, suspension, and school instability rates, with the orange bars representing foster children.

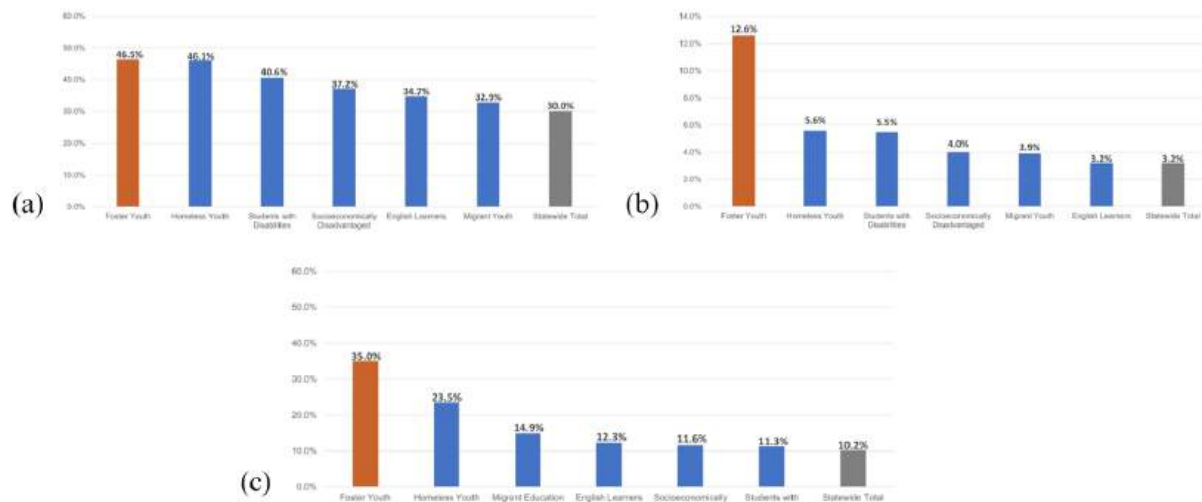


Figure 2. Student education data by group from 2021-22 (a) Chronic absenteeism rates (b) Suspension rates (c) School instability rates [6]

At least one adversarial childhood experience (ACE) is experienced by between one-half and two-thirds of foster children and the effects often extend far into the future. ACEs are more commonly experienced by the Black, Indigenous, and People of Color (BIPOC) population. Therefore, trauma-informed intervention systems are an important resource for these groups. Compared with their peers, foster children are 2.5 to 3.5 times more likely to participate in special education and foster children generally receive a general education diploma (GED) rather than a high school diploma. In determining the diploma of a student, factors like attendance and grades are utilized. Finally, foster children are often placed in congregate care, which increases their placement instability. Both a prolonged stay in congregate settings and an increase in placement changes result in lower graduation rates. Additionally, foster youth often face enrollment and credit transfer issues: a lack of adequate documentation leads to delayed enrollment or foster youth are forced to retake courses, leading to academic setbacks and frustration. The achievement gap is significant, with only 24% of foster students meeting standards in English language arts and 15% in mathematics, compared to 51% and 40% respectively for other students. Addressing these challenges requires collaboration between child welfare agencies, schools, and community partners to ensure foster children receive the support needed to succeed academically and beyond. This paper explores beneficial alternatives and initiatives such as projects and policies to foster school connectedness, increase motivation and

engagement, and raise awareness. Ultimately, this paper hopes to explore the question: What common challenges do foster children face in education, and how can various strategies help address these issues?

Literature Review

Foster children commonly move around to different environments, forcing them to adapt to the circumstance and making it difficult for them to develop a stagnant and belonging identity. Figure 4 outlines this tendency for foster children to change their living placements.

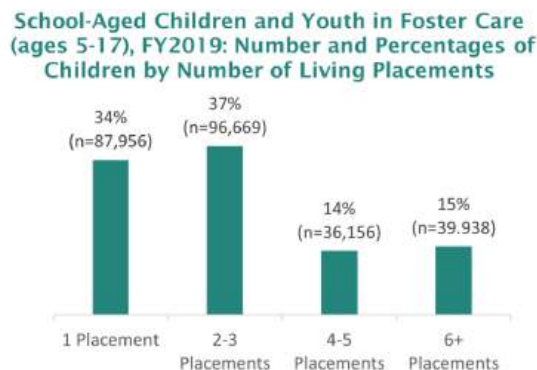


Figure 3. 66% of foster youth aged five to seventeen experience at least one change in placement [4]

On top of this effect on their identity, frequent shifts in placement precipitates various challenges for foster children in education. Zetlin et al. (2010) [7] cites the high rates of absenteeism, disciplinary referrals, grade retention, and depression among foster children. Generally, it has been found that foster children have lower standardized achievement test scores, grades, and poor social skills that often result in aggressive and impulsive behaviors. Through interviews with focus groups, the paper identified various struggles between child welfare agencies (CWA) and schools. On many occasions, foster children were not provided adequate resources in terms of IEP, special education, and retention. For example, one of the foster children did not qualify for special education because they only performed slightly below grade level standards, despite the fact that the student was diagnosed with social anxiety disorder and regularly took antidepressants. As a result, many participants in the focus groups strive for out-of-state programs to receive adequate resources. Additionally, waitlisting and a lack of placement options if a child was determined to be ineligible for special education support were brought up as common issues. Communication was also tagged as a prominent issue. The paper recognized that caregivers often lacked enthusiasm to support their child's education: skipping meetings, unsupportive of school endeavors, and an overall negative orientation towards education. This also meant reluctance to enroll and transport foster youth to after school or weekend programs. Frustration between the school and the CWA was also common: schools believed that CWA's were not actively responding to their concerns and CWA disliked the lack of understanding of the foster care system among educators. Rather than attempting to meet the needs of the foster youth, the paper noted that principals and deans anticipated the removal of

foster youth from their school, primarily because of their tendency to cry or get angry in class. CWA's also pointed at the school's inability to cooperate as schools often regarded their own school policies above national or state-level law. Finally, the paper observed that school liaisons for foster youth prefer to resort to drastic measures such as school transfers or nonpublic school placements because they "don't know what else to do," jeopardizing the stability of foster children.

Somers et al. [3] outlines that residential movements and a lack of coordination between CWA and schools are the main sources of problems such as low educational attainment, high rates of school mobility, and long delays when transferring schools. As a whole, foster children were found to have lower scores on state assessments, miss twice the amount of school, and have an average of a seventh grade reading level as a seventeen or eighteen year old. 65% were found to graduate high school by age twenty-one as opposed to 84% for non-foster children. School transfers mean new classmates, new teachers, new curriculum. Frequent exposure and adaptation to a new environment was found to be associated with lower performance on standardized exams and grades. According to Eckenrode et al. (1995), "15-33% of the negative effects of maltreatment on children's academic achievement was due to high rates of school transfers and residential changes." This effect has also been noticed by social workers in Altshuler (1997) as disruptions such as movement between schools led to declines in performance. In addition to school transfers, paperwork and educational records required for enrollment poses an issue. Powers et al. (2002) found that schools in Pennsylvania reported delays or refusals to enroll foster children because of a need for special education or English language support. This lack of communication between education systems and child welfare agencies also leads to errors in attendance. In some cases, a foster child was never officially discharged from their previous school or marked absent because of visits to the Department of Education. To address this problem, a passport program was initiated. Caseworkers regularly update this record with information on medical, behavioral, psychological and educational status. The Washington's Core Student Record System, is an electronically updated pilot version of the passport program. In the classroom, this lack of communication between the CWA and the school is also prominent. Oftentimes, caseworkers refuse to notify schools of a foster child's status out of fear of stigmatization and teachers are reluctant to take note of this status because of confidentiality. Additionally, teachers and educators often are not well-informed about the system, making it difficult to support the needs of the foster child. The paper also describes potential solutions. To preserve a foster child's school of origin, relaxing policies for residence, providing transportation, and prioritizing placements in the same neighborhood are effective. In New York City, it was found that non-relative foster family placements led to increased school mobility, suggesting the importance of kinship care. For instance, it was found that 26% of children in kinship care, homes where foster children reside with relatives or close family friends, remained in their original community as opposed to 14% of children in non-relative care. Given the abundance of nonrelative placements shown in Figure 5, the commonality of school transfers can be concluded. The primary goal of this paper was to study the role of school connectedness in

influencing a foster child's academic performance. School connectedness was defined as the degree of support in student-teacher and peer relationships, opportunities to be involved in school, and feelings of belonging in this paper. Developmental theories have long cited belonging as an important determinant in psychology and development, influencing features like adolescent risk behavior, mental health, and academic success. In a study conducted in southern China, belonging was a strong herald of satisfaction in school. As foster children experience lower levels of engagement and connection with the school due to their frequent placement changes. While extracurricular activities are effective in increasing engagement, foster children often have limited opportunities to participate. Another intrapersonal factor that was found to play a large role in determining a foster child's academic function is their orientation towards the future. In other words, if they believe in their ability to achieve their goals and have specific aims to achieve, they will feel more connected to the environment. Ultimately, they have a greater likelihood of being academically successful and having pride and positive emotions. This increases motivation, life goal planning, and future occupational outlook, which is correlated with occupational attainment. In terms of external factors, foster children are more likely to drop out of school as their transitions increase. These transitions also decrease their academic growth in reading, writing, and math. The paper found that placement security was a determinant of positive future orientation and school engagement, creating a cyclical relationship. As placement changes increase, internalizing and externalizing behaviors have a greater likelihood of developing, which again increases absences and poor academic achievement. Overall, the paper underscored the importance of school connectedness, which explained variance in academic achievement, and future orientation, which increases school connectedness.

Altshuler [8] analyzes the barriers preventing effective collaboration between public schools and CWA's. Of these barriers, financial considerations, identification of the appropriate population to receive services, and disparate goals were found to be largest. These barriers have been identified to cause poorer outcomes of adult functioning, including a lower likelihood of being employed, maintaining stable housing, and feeling satisfied with their lives than their peers. Additionally, the paper mentioned the importance of academic success, hindered by foster care placements, in determining self-sufficiency and self-esteem. As a result, it is important to address the problems listed in the paper. First, the paper noticed that students and teachers react to foster children. In other words, stigmatization. Referencing negative stereotypes and labels, people tend to have expectations that foster children will have behavioral difficulties. Most of the time, this results in foster children receiving unwanted special treatment. It was also found that foster children often perceive school as an outlet to express their feelings, which are often bottled up at home. As there is no other environment to do so, foster children often channel their frustration and anger at school. Another identified problem was the "opposing" relationship between education providers and child welfare professionals. Due to a lack of trust between the two groups, there is often "fighting back and forth" that mainly stems from a lack of understanding about confidentiality and communication, and little care and commitment. Both sides believe that the other is a main problem, with educators blaming caseworkers for

withholding information and caseworkers believing that educators ask for unnecessary information. Similar to previous findings, educators believed that caseworkers were not actively involved or aware of the foster child's life. Caseworkers concluded that educators do not "want" foster children in their schools and invest minimum energy in supporting their education, largely due to the fact that they anticipate a school transfer in the near future. In terms of solutions, the paper identified needs that should be met from both groups. This includes a sensitivity from educators about the situation of foster children, but refraining from treating them differently. If foster parents are actively involved in their child's education, such as contacting the school and figuring out problems together, foster children were found to succeed better at school. Training for foster parents and caseworkers about education policies and laws was identified to be an important component. The idea of a cross-training, where school personnel shadow a caseworker and vice versa, was also proposed to foster a sense of mutual understanding between the two groups. Supplemental support, such as one-on-one tutoring, peer counseling, social skills training, mentoring, and fun activities are beneficial for at-risk students. Caseworkers should also be aware of these services, and teachers could provide mentoring for foster parents. This would increase communication and allow for the foster parents to convey their concerns. Foster parents would also have a better understanding of the school system, providing them with the information necessary to communicate their needs.

Methodologies

This paper utilized an observational case study and grounded theory approach to study the issues involved in the education system for foster children. By analyzing elements like motivation and school connectedness, data collection was primarily conducted through interviews of individuals experienced in working with foster children's education. This includes employees at the County Office of Education, members of nonprofit organizations, and school personnel that have interacted with foster children. Data was collected through qualitative interviews. Interview questions followed the general structure of the experience of the individual in working with foster children, struggles that they have first-handedly perceived, as well as potential solutions to these issues. The interview questions slightly differed depending on the role of the individual and follow up questions, not listed, were also asked. These questions can be found in the appendix of this paper.

The Otter.ai² application was utilized to audio tape the interviews as a reference for the author. Interviews generally lasted between ten to thirty minutes. The author then worked with nonprofit organizations to address some of the identified issues through new initiatives. Interviews were the primary sources in this paper and the main source of data. Secondary sources were mentioned in rudimentary research of the topic and were used to help create survey questions and ideate initiatives. Through these methods, the paper hoped to identify issues that call for the attention of individuals and potential data analysis, policy, and advocacy projects. All

² <https://otter.ai/>

participants consented to an audio recording and interview prior to the interview process and the questions were approved by the author's mentor.

Results and Findings

A total of nineteen individuals, that can be categorized into three groups, were interviewed: members of the school or district staff, employees at the County Office of Education, and individuals from nonprofit organizations. These interviewees share the mutual characteristic of having directly interacted with foster children in their career. The main ideas from these interviews were extracted and are presented below. Questions asked during these interviews differed based on the interviewee's responsibilities and area of oversight.

School and District Staff

L, a guidance counselor that has four years of experience working with around two to five foster children was also interviewed. As a high school secondary counselor, most foster children L has worked with are in grades nine through twelve. At L's school, prioritizing "consistency and continuity with education" is important, and a foster child was able to remain in their school-of-origin for a short period despite transportation times. While transitioning to a new school is difficult for all students, L noticed that "students that are within the foster care system, they kind of already have some social skills in their toolbelt, so they're able to connect with other students pretty naturally." However, the transition remains difficult from an "academic standpoint and the social standpoint." Additionally, L believes that proximity and feeling connected is a main factor in determining chronic absenteeism: "It's a matter of building community to make sure that they feel comfortable coming here, they've established connections, authentic relationships with teachers and students." and "involving all students within the fabric of the school." In L's experience, documentation issues have not been prevalent due to a prioritization of a foster youth's learning and engagement in their new academic community. Course credit transfers have not been an issue either, as most of the times, courses in the new school are matched with the courses listed in a foster youth's academic records. L also cited the importance of helping foster children understand the available embedded supports, especially for schools that vary in academic rigor. In the past, L has noticed a high motivation in a foster child's desire to learn: "They take school very seriously...they're controlling the controllables...and because of the soft skills they've established through their lives, they're high on self-advocacy." The main cause of this is that a foster youth's academic performance is something that they can control: "Things can be moving around outside of the academic institution wall...that they don't have control over. So the things that they do have control over, they lean into it." During the interview, L emphasized the need to be cautious about the variance in comfort level of a foster child: L believes that for a foster child to unpack their story, the conversation should happen organically: "If students build connections, and they build friendships...then those stories are expressed organically." Using a backup metaphor, it is important to be sensitive about the diversity of a foster youth's story: "Everyone has a story, and

when we are unpacking that story, it's not just packing and putting it on the table. It's also taking those pieces...and putting it back in their personal backpacks that they carry every day." In terms of general challenges, systems is the first thing that comes to L's mind. The constant need to "weave through" different institutions and repeatedly learn and unlearn them. As a whole, L underscores the importance of empathy in individuals that work with foster children:

People come from different backgrounds...[and] different upbringings. The one great equalizer is education. So to be able to meet students where they are, no matter their background, no matter their upbringing. We need to understand that everyone's story is unique and...[that] makes them who they are...If we have an understanding of that, then we're able to meet students where they are and provide them the avenues and resources to find success academically. Not only in a K-12 setting, but also preparing them for higher education and to continue to strive and reach their goals.

County Office of Education Employees

Employees from the Santa Clara and San Mateo County Office of Education were contacted and interviewed in this paper. The first individual interviewed from the San Mateo County Office of Education (SMCOE) was an educational liaison that will be referred to as L. L has two decades of experience working with foster children and currently specializes in working with senior year high school foster youth students and AB Non-Minor Dependent (NMD) foster youth. NMD youth stay in the foster care system until the age of twenty-one and directly receive money. In the past, L has also worked with high school students of all grades and K-12 students. In terms of school instability, L remembers working with a high school senior that moved four times within the months of August and December. In terms of challenges L has perceived throughout L's career, the main challenges have been making friends, staying in a child's original school, and court. Oftentimes, foster youth miss school because of court appearances or visitations that generally take place during school hours. AB 490 labels court and medical visits as an excused absence for foster youth. Additionally, L has found it difficult for foster children to enroll in school if they have an IEP. This is primarily due to the long process of approval: a program specialist is required to review the IEPs and ensure a comparable offer can be offered in this new school, a process that often takes two to three weeks. Most of the time, foster children are enrolled within two business days in their new school but cannot attend and receive their IEP until the documents are reviewed and a meeting is held. Missing school for around three weeks because of IEP placements are significant disruptions in a child's education. While partial credit transfer is a right, it is "not the letter of the law" and so foster children often need to advocate for themselves. As a result, L mentioned that it would be helpful to compile the laws and practices for foster children to understand what is written in the law. L observed that high school students are more vocal about advocating for their own rights while middle school students are less likely to advocate. As a whole, L described that foster children are "parentified at age eight" and recognize the need to care for themselves early on. In terms of training for educators, L

recommended the YouTube movie *ReMoved*³, which is already used for social workers during their training. This film, which covers the foster child experience and the basics of the system, can increase empathy and understanding for teachers and other educators that interact with foster children. L noticed that “foster children love school” because it is “one place they can be a kid” and can escape the societal pressure to be parentified. As opposed to other places, which could be associated with traumatic experiences such as sexual abuse, school is a fresh start. L recognized that in terms of motivation, foster children are eager to be at school, but “they don’t know how to be there safely.” As a result, L suggested the idea of a welcoming basket or all-star club in collaboration with a foster child’s school ASB. The welcoming basket could be initiated as a state-wide recommendation, where any student that is eligible to receive free school lunch through the Title I program will receive a basket. In this basket, L envisions anything along the lines of a letter from the ASB president, a water bottle, a sweatshirt, tickets for a game, a prime access code, a free yearbook, or other fun activities “that makes them feel welcome.” In order to combat stigmatization issues, L suggested to refrain from mentioning a foster child’s status, and instead give out baskets through a student’s counselor. Many times, L has observed that foster children choose to voluntarily skip school out of embarrassment: not having money for dress-ups, basketball game tickets, and being “ashamed because they can’t participate.” L identified this issue as a possible platform of collaboration with a school’s Parent Teacher Association (PTA). Through the all-star club, an opt-in initiative, ASB students interact with foster youth. If foster youth are comfortable sharing their status, ASB students will know of their status but not their full situation. This provides the opportunity for foster children to make friends, as “belonging is key.” Over the years, L has noticed a tendency for foster children to “give up” because of the large workload they are required to make up in order to attend college. With a peer mentor, some of the frustration can be eliminated and the foster youth have a source of support.

D, the Foster Youth Services Manager from the Santa Clara County Office of Education that has worked with foster youth for around two decades, was also interviewed. D has worked with foster youth of different ages, and currently works with around 120 caseloads of transitional kindergarten (TK) and post-secondary students. When D was a teacher, many of the current laws were yet to be enforced. As a result, D was not aware of the statuses of the students in D’s classroom and whether or not they had approved IEPs. While the law says that teachers should know of a foster youth’s status, each school district works with different regulations. In some districts, this information is not disclosed to teachers, which can be detrimental. D has noticed that schools that notify their teachers of a foster youth’s status often have kids that “do much better.” However, this is not due to special treatment from the teacher. Instead, “it’s about having a little bit more sensitivity and awareness. Being trauma-informed is the key.” For foster youth, D has observed that there are often gaps in their learning and therefore they need an extra push. D outlines the importance of training not just teachers and counselors, but also custodians, cafeteria staff, janitors, and other staff should be aware of trauma-informed practices. Many

³ <https://youtu.be/IOeQUwdAjE0?feature=shared>

times, foster children form intimate relationships with these individuals and view them as role models. D sees the need for trauma-informed training to be established as a national policy: “It’s not [a foster child’s] fault that they’re in the system. How can we continue to punish them by not acknowledging what they’ve been through?” As foster youth might not “have the capacity to react the way that you would expect them to react,” D says it is easy to initiate a “negative response because you don’t understand what you’re seeing.” Without a full understanding of what foster children have been through, D has observed that responses are often that foster children are “lazy and not motivated.” D remembers a personal experience that resonates with this.

“There was a kiddo that was yelling, and I was new at the school. I didn’t know him. I just yelled at him, thinking, ‘Oh, just [a kid needing] some attention.’ [I find out later that] this kid had witnessed a murder the night before. He had slept out on the street, and he was basically at school just to get a meal and feel safe.”

In D’s opinion, the main reason that trauma-informed training has yet to be normalized is because of the traditional barriers of institutional racism and a lack of urgency. D believes there needs to be more prioritization of policies that address issues like these. At the County Office, records are requested for all students under their caseload to monitor progress. Of the foster youth under D’s caseload, three are refusing to attend school. Two of them have experienced two or three different home placements within the last school year and were placed out-of-county. Additionally, most of the foster youth are placed with Caucasian families that do not share their culture. While schools may try to engage the foster youth, most of the time, D has noticed that “they haven’t quite figured out what they need.” For one of the girls under D’s caseload, she is being pushed into gangs and dealing with illegal actions. She also has an IEP as it is difficult to keep up academically in a traditional education setting. D explains that, for her, “[She doesn’t] want to go to school because [she’s] embarrassed. [She’s] behind, [and she] feel[s] stupid.” Despite the resources she has access to, such as a social worker, CASA, and more, the true deciding factor in whether or not she attends school is if she makes up her mind. “She can’t hear me until she’s ready to accept that, you know what? I need education.” In the past, she was involved in a fight and the school took that chance to suspend her, even putting her up for expulsion. D had to fight for her rights in order to keep her in school, and found that the school was more proactive in implementing intervention and support programs once D advocated for her rights. When asked about the school’s response, D believes that the difference between schools and a lack of proper training in foster youth rights account for their decisions. “For years, it’s all been about remov[ing] the child, because that’s going to change the behavior. No, it’s not. Not for foster [children].” Instead of making assumptions, D believes that educators need to understand the reason behind why a foster youth may be “acting out.” Having a structured and dependable support system is also important: for instance, D says that “she doesn’t like where she’s living. She can’t go back to her mom, and she doesn’t want to go to the hub.” In terms of motivation, D has observed it to be “all over the place.” Many times, D has realized that foster

youth are refusing mental health services, for example, but later learn that they do need access to the provided programs. In Santa Clara County, foster youth are able to return and receive various resources if they need to. One of them under D's care, currently has a 4.0 GPA and knows she wants to be a doctor.

R is an educational liaison at SMCOE and has worked in this position for ten years and has worked with children in long-term foster care as a social worker for six to eight years. Throughout R's career, R has worked with hundreds of foster children of all school levels and ages, but currently works with foster children in kindergarten through eleventh grade. In R's experience, transitions have happened almost "every day," except during COVID, when an online format was used. However, this online format was "detrimental" in the long term but could be explored as a temporary education setting when foster youth are unable to attend school. The transition comes with a "loss of friends and trusted adults," and adapting to the new environment also means "new rules [and] new teachers." As a whole, foster youth experience trouble building trusted relationships and their frequent changes in school causes them to be overrepresented in disciplinary action. Oftentimes, the teacher does not know the foster youth well enough in order to make accurate judgements about their behavior. While SMCOE has provided training, it is difficult to reach teachers because there are few professional development days and there is already an abundance in required training for teachers. For SMCOE, there are twenty-three districts under the county, but not all of them have students involved in the foster care system. As a result, it is difficult to mandate training. As a foster youth's status is private information, it may be helpful to create an opt-in or opt-out system for foster youth to determine whether or not a teacher is aware of their status. At the same time, this can provide the opportunity for a foster youth to share their preferences in the classroom. R has also frequently observed chronic absenteeism, documentation issues, and course credits not transferring, but has noticed that access to enrollment history helps with course credits. In terms of motivation, R says that foster youth often have a "why bother?" mindset due to their frequent movement from school to school. Overall, R sees the education gap, even without disruptions to education, as one of the largest challenges foster youth face. On top of that, schools have vastly different curriculums which makes learning much more difficult. School connectedness is also important, as one parent in Carlmont High School started a program to provide care packages for foster youth and homeless students. Additionally, having foster youth initial meetings was found to benefit the school in helping them understand a student's situation. This is a suggested practice to better understand the student, and R also recommended the practice of providing academic assessments to pinpoint the educational gap. Resources would then be provided based on these results to help the foster youth feel connected and supported as soon as possible.

Individuals from Nonprofit Organizations

The Academic Services Manager M from Foster Love⁴, who was previously an educator and has worked with Foster Love since 2019, was interviewed. M generally works with

⁴ <https://fosterlove.com/>

transition-aged youth from sixteen to twenty-four years of age, although the average age that M has worked with is eighteen to nineteen. Primarily, M helps foster students apply for scholarships and receive support throughout their academic journey. In M's experience, 60-70% of foster children experienced large school changes that caused a disruption. Additionally, M noticed that foster parents have an important role of motivating foster students to learn. One example of this was shared: a student that frequently moved families and entered the system at a young age received support from his foster parents in his junior year of high school. These parents "convinced him to focus on academics" which "changed his view on everything." He graduated with a 4.0 GPA and is currently attending Amherst College. As foster children "see a lot of faces in the system, change social workers, judges, lawyers, and police officers, etc." there is a lack of consistency. As for a reluctance to form relationships at school, M notices that the simplest solution, keeping foster children in the same school, is not always feasible. The government can compensate a foster parent for transportation costs to this same school, but fostering school connectedness, which can be achieved through participation in sports and clubs, is just as important. Funding for these extracurricular sports and clubs unaffiliated to a certain school is also important. As foster children frequently experience school changes, participating in one school's club is fungible. M also mentioned an important point: stability at home can be perceived as a prerequisite for a motivation to learn and participate at school. If school becomes an "outlet and safe space" to "escape what's happening at home," home becomes a "negative community" that pushes foster children "down." Effectively, going home diminishes any sparks of motivation or feelings of positivity the children might have experienced at school. M believes a lack of stability or comfort at home means that foster children have "no gasoline" to power themselves through school, as home is not an "open environment for them to share their motivations." In terms of challenges, M mentioned a few that they have witnessed in their career: not having enough food at home, foster parents disliking other children in the home, a lack of transportation for events like prom, and a lack of access to technology. For technology, M noticed that foster children often received old or reused computers, and often do not have access to internet and printers. With the increase in technology use at school, this renders it difficult for foster children to continue their studies at home. As a whole, M believes that foster children "need consistency" and support for "everything from A to Z," such as career goals and technology. Identifying the end goal and creating a plan to achieve this is important, yet schools often do not adequately address this issue. Many times, resources are cut off at a certain age, which is restrictive for students that have yet to graduate high school or complete their studies. As a result, a policy change to extend the age range or data analysis of the benefits of certain age caps was something that M found would be worthwhile. Overall, M recognizes that California is well-ahead in the foster children realm. While California has implemented many changes, there is a lack of government support outside the state. There is a need for other states to catch up to California's pace, as California is a "leader in the foster care movement."

D, a member of the California Foster Youth Education Task Force and is an Associate Director at Children Now⁵, specializing in child welfare and foster care policy. When asked about mandated training for educators, D shared that the “turnover” involved with the foster youth liaison position means that “even the person who’s supposed to be the expert doesn’t always know everything they’re supposed to be doing.” Additionally, the abundance of moving pieces within the scheme of school adds to the confusion. D suggested connecting with the County Office of Education and identifying the current requirements before designing a new training program. However, D warned about a mandated training system: “Anything that’s mandated for educators to do will get a lot of pushback, because schools are overwhelmed with all the stuff they have to do.” This results in frustration for advocates, so D suggested expanding on the current system by training other groups before mandating the system through a bill. While educators receive trauma-informed training, they often do not have enough understanding to look at the issue from a foster child’s perspective: “For example, high rates of suspension among youth in foster care...[Educators] should have some understanding that students of various backgrounds, not even necessarily youth in foster care...are going to be dealing with trauma, and their reactions could be symptom of that. Things like suspensions, taking them out of learning, isn’t going to solve that.” D also noted that looking at the differences between counties would be helpful. As for funding extracurricular activities unaffiliated with schools, D identified that since foster children only change schools if the distance between their new and previous home is large enough, an unaffiliated club would likely be too far from their new school. While a foster child’s right to their school of origin is fundamental, D has noticed that there are often challenges involved with implementing that: “If the caregiver...doesn’t know to [ask about a foster child’s school of origin], they just enroll them in the school that’s closest, and [foster children] lose that connection.” D also sees data analysis on the differences in age caps feasible. However, D cites that expanding these programs and the groups that it serves is difficult from the funding perspective. Through D’s own research, it was found that through surveys and focus groups, foster children believe the ILP should be raised to at least age twenty-four. However, when this bill was proposed a few years ago, funding became an issue that prevented it from passing. Since the ILP is backed by federal funding, California receives a certain amount that is then distributed to the counties. The “counties have a lot of leeway on how they use the money and what types of services they provide,” which means that “there’s already some counties that serve older youth over age twenty-one.” Ultimately, the bill was not passed because “all the counties have to opt into something, or none of them can [expand funding].” Expanding age caps remains a priority as the services offered include career assistance, something that foster children “don’t really figure out...until after [age] twenty-one and then in those few years later, they’re like ‘Oh, now I need this help. I didn’t need it before, but now I do.’” However, D believes that data analysis regarding age caps in foster care programs would still be beneficial: “If you did some analysis off, ‘here’s the various programs, and here’s the age they end at, because a lot of folks are trying to get more consistency...or former foster youth know[ing] what services are going to shut off

⁵ <https://www.childrennow.org/>

when.” These programs, such as extended foster care or transitional housing programs, often differ across states and even counties.

Additionally, various initiatives are currently being planned based on the responses and input received from the interviews. First, a writing competition for foster youth held in May, which is Foster Care Awareness Month. This initiative is intended to boost motivation and academic curiosity, while providing a platform for foster children to explore creative writing. Through the author’s own experiences with creative writing and other sources, creative writing can act as an outlet of expression and an opportunity for foster children to share their stories. The participants in the competition will also have the opportunity to interact with each other, creating a safe community to make friends and share interests. Ultimately, the core idea of the writing competition is for foster children to partake in effective academic writing. As a result, the competition intends to invite credible judges to select winners for the competition and provide feedback, if possible, to the participants. These reviewers will receive a small stipend for their contributions and the winners of the writing competition will also receive a monetary award. The author hopes to receive funding from the Palo Alto ThinkFund for these purposes. While there are no age or location restrictions to participation, winners will likely be categorized by age group and if possible, an in-person gathering will be held for the participants. If not, the author intends to host a virtual meet-up to foster a sense of community. Pivotal, an organization focusing on career development for foster children, will collaborate to host this competition and other nonprofit organizations and County Offices of Education will help with recruitment.

Next, the author is communicating with the school guidance counselor and County Office of Education to establish an ASB welcome basket program for foster children. This package would include the basic items that students receive during freshman orientation (a school water bottle, campus map, planner, etc.) along with other items such as school merchandise and gift card incentives from local stores. While this initiative intends to start at the high school and middle school level as a part of the ASB programs, the author hopes to expand to elementary schools in the future. The package will also include a letter from an ASB member, where their name can be included at their discretion, but personal information about the foster children is not provided. Students will not have access to the information of any of the foster children on campus and generic packages will be included in the counselor’s office for distribution when necessary. This will likely be an opt-in program, where a foster child can decide whether or not they are comfortable and would like to receive this welcome package. The SCCOE can provide funding for some of the items included in the package.

Additionally, the author hopes to increase accessibility to foster care training systems for educators. This includes counselors, teachers, and other staff that come into contact with foster children in their work but are not required to complete training. By doing so, the author hopes that these educators will better understand the foster child experience and thus meet their needs better. While existing training systems exist, it may be helpful to look into mandating these trainings. Whether at the school, city, district, or state-level, providing more information and knowledge would be beneficial. In order to analyze the impact and effects of a certain age cap

for foster care, data analysis was also conducted. By analyzing variables such as placement stability rating and foster parent support rating, correlation was tested through Python. The results of the data analysis are shown below, and included all fifty states. The compiled data was based off of the State-by-State Age-Out Database for Children in Foster Care⁶ and Child Welfare and Foster Care Statistics from the Annie E. Casey Foundation⁷. First, the correlation heatmap between the variables was produced, which can be seen in Figure 6. The variables, their possible values, and the utilized dataset are listed in the table below.

Analyzed factor	Possible value range
Foster Care Age Cap	18-23, age value
Placement Stability Rating	1.0-5.0, rating
Behavioral Health Services	1, 0: yes or no
Adulthood Transition Support	1-5, rating
Foster Parent Support Rating	1.0-5.0, rating
Eligibility for Extended Foster Care	1, 0: yes or no

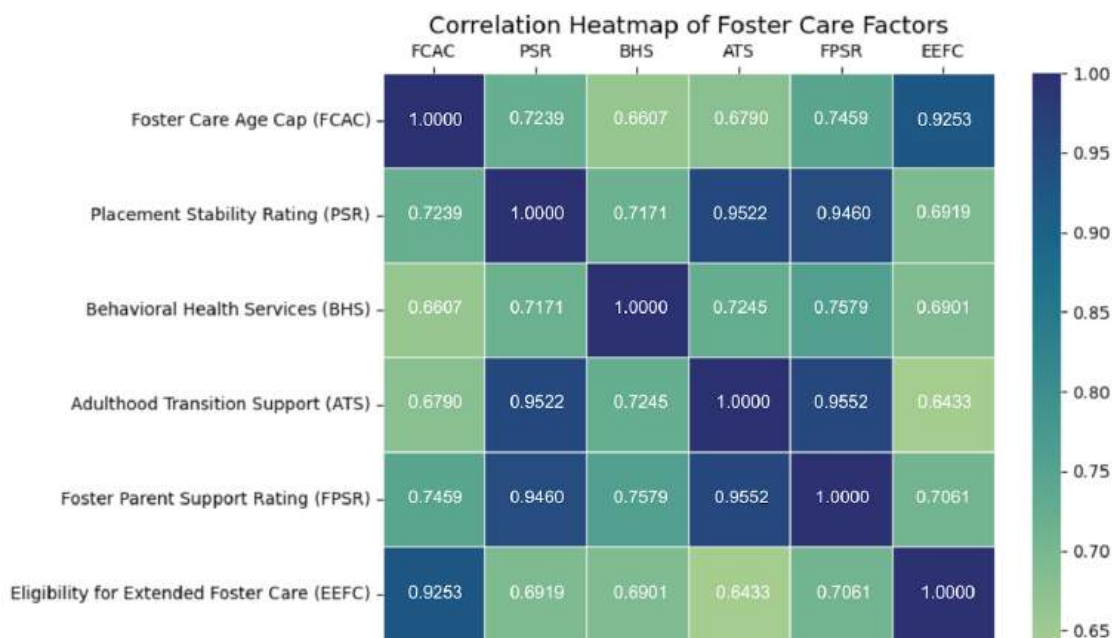


Figure 6. Correlation Heatmap of Analyzed Factors

⁶<https://projectlifeline.us/predicate-order-resource-center/state-by-state-age-out-database-for-children-in-foster-care/>

⁷ <https://www.aecf.org/blog/child-welfare-and-foster-care-statistics>

Each grid box corresponds to the correlation coefficient of the two variables accounted for. The color differs based on the strength of the correlation, with a dark blue correlating to a 1.0 and a lime green correlating to a 0.5. This provides an easy visual that assists with the conclusion that can be reached through this data. From the heatmap, it can be seen that behavioral health services and adulthood transition support had the lowest correlation with foster care age cap, the focal point of the analysis. Similarly, eligibility for extended foster care had the highest correlation. As all of the correlation coefficients are positive, the relationship is also positive. For instance, as the foster care age cap increases in a state, there is a greater likelihood of having eligibility for extended foster care.

Conclusions and Next Steps

Through interviews with individuals that have directly interacted with foster youth, common and local challenges encountered by foster youth were investigated. From school and district staff, the most common issues faced are shown below with the number of staff that mentioned each issue. Overall, these findings indicate that the challenges described below have various effects. A lack of school stability can affect emotional wellbeing, academic achievement, and a foster youth's social skills. Without a standardized state-wide or nation-wide course credit transfer process, schools often have to adjust to meet course credit transfers. If adjusting fails, foster youth are forced to repeat their credits. Proactive programs and continuity across schools helps with school engagement and feelings of belonging, although these standards differ based on a foster youth's personality and extrovert character. Foster families that are supportive of their child's education are often observed to positively impact students, increasing their academic and social success. Due to frequent transitions and past experiences, foster youth may feel less comfortable forming intimate relationships. They may develop an avoidance towards these relationships, fearing the possibility of future situations that may cause them distress. Staff that directly interact with foster youth should also be well-trained and aware of the unique needs of foster youth. Additionally, better communication between different agencies and entities would benefit foster youth in their educational journey.

Challenge	Number of staff that mentioned the challenge
School instability	3
Course credit transfer	2
Low feelings of school connectedness	3
Support from foster family	2
Chronic absenteeism	2
IEP or academic accommodation approval delay	1
Comfort in sharing foster youth status	3
Institutional barriers in the system	3

For the seven interviewed individuals from the County Office of Education, the same table was constructed. School changes were identified as a significant disruption in an academic and social aspect. As a result, foster youth often require more attention and an extra “push” to make up for their gaps in learning. With these changes comes a difficulty in enrollment, especially in IEP and other accommodation services, due to approval processes and required documents. School staff generally lack training in trauma-informed practices and an awareness of how the foster care system works. To prevent misunderstandings and potential triggering behavior, it is important for teachers to not only be aware of a foster youth’s status, but also be well-trained before interacting with them. This sensitivity should also surface in the form of special treatment, as individuals emphasized the need to stray away from any information that could remind a student of their foster care status. Ultimately, this translates into no stigmatization, special treatment, or discussion about the system without a child’s consent. Home placements where foster youth struggle with bicultural disconnect was also mentioned to be detrimental and an additional challenge. Finally, it is important for foster youth to advocate for their own rights. While a foster youth bill of rights exists, not every child is well-aware of what that entails. Therefore, increasing accessibility and awareness of these rights would allow foster youth to be more vocal about their own rights, especially for younger children.

Challenge	Number of staff that mentioned the challenge
School instability	5
Enrollment and IEP difficulty	1
Missing school and absenteeism	1
Emotional and behavioral readiness	2
Difficulty in forming meaningful connections	4
Academic discrepancies	2
Lack of educational motivation and awareness	1
Placements with mismatched cultures	1
Stigmatization	1
Lack of trauma-informed practices	3

Six individuals were interviewed from nonprofit organizations and the table is shown below. The mentioned challenges suggest that frequent school changes are detrimental to a foster youth’s education and connections. While connecting and socializing is important, foster youth often struggle to form meaningful relationships due to their school changes and comfort level. An online education format would not solve the problem due to the lack of socialization and

increased feeling of isolation. Without consistent support across schools, such as academic and emotional, foster youth often fall behind. More access to resources like transportation to a school of origin, technology, or opportunities to engage in extracurricular activities would enhance feelings of school connectedness and improve achievement gaps. With the increase in technology and internet use for school assignments, these resources are crucial to helping a foster youth excel in education. A stable and supportive home environment encourages foster youth to engage in education. On the other hand, instability and negative influence at home can undermine a foster youth's motivation and their efforts. An increase in trauma-informed practices would ensure that teachers are effectively meeting the needs of foster youth. With heightened awareness, teachers would understand and look at the situation from a foster youth's perspective, helping form relationships more easily and foster youth's willingness to share their worries and show up to school. Between schools, different policies exist: for instance, some schools may not implement certain rights or procedures. For many different programs, the age limit varies between states and counties. While it would be helpful to extend these programs beyond the age cap, many of them being age twenty-one, funding is not sufficient to support this. This is important as foster youth often only discover the importance of education and start to consider their future at this age. Therefore, it is important to advocate for funding and the extension of age limits to support foster youth in their future careers and goals.

Challenges	Number of staff that mentioned the challenge
School instability	4
Low feelings of school connectedness	4
Consistency across schools and systems	1
Need for extracurriculars and transportation support	2
Lack of technology access	1
Home instability decreasing motivation to learn	1
Lack of trauma-informed practices	1
Lack of awareness in school staff	1
Inefficient support services	1
Insufficient communication between foster youth liaisons	1
School of origin implementation	1
Lack of funding to sustain programs	1
Career assistance beyond age caps	1

The information concluded shows that various challenges exist for foster youth in education. On a broader scale, many of the challenges faced by foster youth are institutional and systemic barriers. These roadblocks have translated into different academic and social challenges: from achievement gaps and a lack of willingness to attend school, to an aversion towards intimate relationships with peers and adults alike. The paper notably investigates some of the issues present at the local level in accordance with the interview responses from individuals experienced with the foster care system. Overall, school connectedness, a supportive family and stable home environment, and need for well-informed staff members is important. In terms of frequency, the most commonly mentioned barriers were school instability, feelings of school connectedness, a difficulty in forming connections and relationships, and a lack of trauma-informed practices. While there are many other issues across the board, an abundance of these factors are linked hand-in-hand. For instance, more consistency across schools could positively impact communication between liaisons and the rate of school of origin implementation. With more funding for programs, support systems can be more effective in supporting the needs of foster youth. More access to technology would increase a foster youth's motivation to learn, and placements with families with a mutual culture would create a safe and welcoming environment for foster youth to excel in academics. Challenges can vary based on location and an individual's unique experiences, and this paper explores some of them. Due to the lack of research in foster care, it is important to continue to conduct relevant research in the field. In the future, interviews with foster youth, their parents, their teachers, and other individuals that directly work with foster youth would be recommended. Restrictions from the IRA due to the author's role as a high school student hindered some of the interviews mentioned above. However, it is crucial for future interviews with populations directly impacted and involved by the foster care system to be sensitive and considerate towards the interviewees. Additionally, more data analysis, such as the correlational heat map generated in this paper, and data-driven policy or bill initiatives would help in establishing effective and targeted solutions.

Works Cited

1. Chipungu, Sandra Stukes, and Tricia B. Bent-Goodley. "Meeting the Challenges of Contemporary Foster Care." *The Future of Children*, vol. 14, no. 1, 2004, pp. 75–93. JSTOR, <https://doi.org/10.2307/1602755>. Accessed 8 Oct. 2024.
2. Mieke Denuwelaere, and Piet Bracke. "Support and Conflict in the Foster Family and Children's Well-Being: A Comparison between Foster and Birth Children." *Family Relations*, vol. 56, no. 1, 2007, pp. 67–79. JSTOR, <http://www.jstor.org/stable/4541648>. Accessed 27 Nov. 2024.
3. Somers, Cheryl L et al. "Academic Achievement Among a Sample of Youth in Foster Care: The Role of School Connectedness." *Psychology in the schools* vol. 57,12 (2020): 1845-1863. doi:10.1002/pits.22433
4. Legal Center for Foster Care and Education. *Foster Care & Education Fast Facts*. American Bar Association Center on Children and the Law, Jan. 2022, www.fostercareandeducation.org. Accessed 20 February 2025.
5. ABA Center on Children and the Law. *Exploring Education Outcomes*. January 2022, <https://static1.squarespace.com/static/63dcf65b8d0c56709027332e/t/6454f86f0f4c483811518c22/1683290224435/Education+Outcomes+final.pdf>
6. Wondra, Danielle, et al. *Supporting School Attendance for Students in Foster Care*. Children Now, Sept. 2023, www.childrennow.org. Accessed 20 February 2025.
7. Zetlin, Andrea, et al. "Caregivers, School Liaisons, and Agency Advocates Speak Out about the Educational Needs of Children and Youths in Foster Care." *Social Work*, vol. 55, no. 3, 2010, pp. 245–54. JSTOR, <http://www.jstor.org/stable/23719080>. Accessed 10 Feb. 2025.
8. Altshuler, Sandra J. "From Barriers to Successful Collaboration: Public Schools and Child Welfare Working Together." *Social Work*, vol. 48, no. 1, 2003, pp. 52–63. JSTOR, <http://www.jstor.org/stable/23718797>. Accessed 20 Feb. 2025.

Exploring Wax Blends with UV Stabilizers and Antioxidants for Enhanced Solar Panel Applications By Kirti Ponnaganti

This study investigates the development of advanced wax formulations designed to enhance the durability and efficiency of solar panels under diverse environmental conditions. Given the critical role of protective coatings in prolonging the lifespan of solar panels, the properties of beeswax and its blends with carnauba wax in varying ratios (5:1, 4:1, 3:1) were studied to create a natural form of coating to apply on solar panels. These ratios were selected to evaluate how different compositions of waxes affect key performance indicators, including melting points, thermal stability, and resistance to degradation, which are vital for withstanding harsh conditions. To further improve the resilience of these wax formulations when applied on solar panels, antioxidants like Vitamin E and UV stabilizers were incorporated. This addition aimed to mitigate thermal aging, a common issue faced by solar panel coatings. Accelerated aging tests were then conducted to simulate long-term performance and assess the efficacy of the enhanced formulations. Ultimately, this research seeks to optimize protective coatings for solar panels, thereby contributing to improved efficiency and longevity in renewable energy systems. The results indicate that waxes with high melting points and thermal stability, particularly beeswax, significantly improved solar panel efficiency. Additionally, the inclusion of carnauba wax further enhanced the performance of the formulations, contributing to greater overall durability and effectiveness in renewable energy systems such as solar panels.

Introduction

With the growing demand for renewable energy sources, optimizing solar panel efficiency has become a critical focus in sustainable technology development. Solar panels face numerous environmental challenges, including dust accumulation, UV radiation, and temperature fluctuations. These issues significantly reduce the lifespan and energy output of solar panels, a critical concern as the energy sector increasingly shifts toward solar energy as a renewable solution. (Jampole and Chaudhary 3-4)

The dust and dirt accumulation on the surface of solar panels blocks sunlight, leading to drops in energy conversion efficiency. This problem is severe in arid regions or areas with high pollution, a phenomenon explained by Fresnel's laws of reflection and refraction. (Xie et al. 24-25) Prolonged exposure to UV radiation causes materials in the panels to degrade, reducing their efficiency and lifespan. Additionally, thermal stress from high temperatures causes microcracks and damage to photovoltaic cells within the solar panels, impairing overall performance. These issues hinder the effectiveness and sustainability of solar panels (Masterson 22).

Current solutions for addressing the challenges faced by solar panels include mechanical cleaning methods, water-based cleaning systems, and synthetic protective coatings. However, mechanical cleaning, which often involves manual or robotic systems, is labor-intensive and costly to maintain. Moreover, repeated physical contact leads to micro-scratches on a panel's

surface, reducing efficiency. Additionally, water-based cleaning systems are commonly used but pose challenges in water-scarce regions, especially as the demand for solar energy rises. Synthetic coatings on solar panels improve efficiency by reducing dust and enhancing light transmission but can be costly, degrade over time, and require frequent maintenance (Aljaghoub 1-3). These existing approaches, while offering some benefits, fail to provide a cost-effective, durable, and sustainable solution to challenges faced by solar panels.



Above is a picture of a solar panel heavily affected by dust and dirt accumulation.

One promising approach is the application of wax coatings to solar panel surfaces. Previous research has highlighted the reflective and protective properties of certain waxes, which can mitigate heat buildup by reflecting excess sunlight while maintaining light transmission to photovoltaic cells. Natural waxes such as carnauba and beeswax have demonstrated high thermal stability and reflectivity, making them strong candidates for this application ((Turner 2). Synthetic additives, including antioxidants and UV stabilizers, have shown potential to further improve wax formulations by enhancing resistance to environmental degradation.

Despite these promising findings, limited research has focused on optimizing wax formulations specifically for solar panel protection. This study addresses this gap by exploring the properties of natural and synthetic waxes and developing blends tailored for solar panel applications. By manipulating wax properties through blending and incorporating additives, the research seeks to improve their thermal stability, UV resistance, and overall durability.

The objectives of this study are to evaluate the melting points, thermal stability, and degradation resistance of wax blends, with a particular focus on enhancing solar panel performance. It is hypothesized that formulations containing a higher proportion of carnauba wax, combined with antioxidants and UV stabilizers, will provide superior protection by reducing heat-induced degradation, enhancing UV resistance, and minimizing surface damage.

Materials and Methods

For this study, several materials were used to prepare the wax blends and test their properties. Carnauba wax and beeswax served as primary components for wax blends. The UV stabilizer and Vitamin E powder were added to was blends to enhance UV resistance and antioxidative properties. Mineral oil was used to create an oil bath for thermal testing of the wax blends. Micro solar panels were employed to test the performance of the wax coatings under

sunlight exposure. Voltage readings were taken using a multimeter, and temperature was monitored using a thermometer. The waxes were melted and mixed in stainless steel bowls, and small amounts of each wax blend were transferred into test tubes using capillary tubes. After pouring the molten wax onto aluminum foil, the samples were cooled and solidified. A cooking oven-safe tray was used for the accelerated aging test in an oven set to low temperatures. Other tools included a spatula, soft cloth for cleaning the solar panels, and protective gear such as heat-resistant gloves, sunglasses, and safety glasses.

The first phase of this experiment will include the preparation of waxes and thermal testing of specific properties they hold. To prepare the wax blends, beeswax and carnauba wax were measured and mixed in ratios of 5:1, 4:1, and 3:1, with the total weight for each blend set to 10 grams. The waxes were melted over a hotplate, and once liquefied, each blend was poured into separate test tubes and left to solidify. Once the waxes had hardened, small amounts were shaved or scraped and inserted into capillary tubes. The samples were then subjected to a mineral oil bath in a heat-resistant container, where the oil was gradually heated while monitoring the temperature with a thermometer. The waxes were observed for melting points and thermal stability, with the temperature at which each wax began to melt being recorded. The stability of each blend at higher temperatures was assessed by noting any decomposition, texture, or color changes.

For the second phase, a combination of beeswax and carnauba wax was melted in a heat-resistant beaker on a hot plate. Once the mixture was fully melted, Vitamin E powder and the UV stabilizer were added to separate samples. Each formulation was stirred thoroughly to ensure even distribution of additives. Afterward, the wax mixtures were poured onto aluminum foil to cool and solidify. To simulate long-term aging, the wax samples were placed in a preheated oven set to 50–70°C (120–160°F) for several hours. The samples were inspected every 1–2 hours for changes such as cracking, discoloration, or softening. These observations were documented to assess the performance of the wax formulations under thermal stress.

The solar panel surfaces were cleaned using a soft, dry cloth to remove any dust or particles. Wax samples were gently heated to soften them slightly before being applied to the panels using a spatula to ensure an even coating. Each panel was left to cool and solidify before being placed outside for sunlight exposure. The panels were exposed to direct sunlight for a full day to simulate prolonged exposure to environmental conditions. Every hour, a multimeter was used to measure and record the voltage generated by each wax-coated panel section, as well as solar panels with no wax. Observations were made throughout the testing period to document any changes in the wax coatings, such as melting, cracking, or discoloration, and these were compared with power measurements to evaluate the effect of the wax coatings on solar panel performance.

The data collected from the melting point, density, accelerated aging, and solar panel testing will together provide a comprehensive understanding of the wax formulations' properties and performance.

The melting points recorded for each wax blend will offer insight into the thermal stability of the formulations. A higher melting point generally indicates a formulation that can withstand higher temperatures without degrading or melting, which is particularly important for solar panel coatings that are exposed to intense sunlight. By comparing the melting points across different ratios of carnauba wax to beeswax, it is possible to determine how each component affects the formulation's heat resistance. Carnauba wax, known for its high melting point, may contribute to greater stability, while beeswax might lower it.

The accelerated aging test data will reveal each wax formulation's resistance to thermal and UV degradation over time. Visible changes, such as cracking, discoloration, or a change in consistency, are signs of degradation that reduce the wax's effectiveness as a protective coating. Blends that maintain their physical integrity and show minimal changes under high-temperature conditions will be more suitable for use on solar panels, as they suggest that the wax can endure prolonged sunlight exposure ((TeKrony 2). If formulations with antioxidants or UV stabilizers show less degradation compared to those without, it will confirm the effectiveness of these additives in enhancing the wax's durability and lifespan. This insight is essential for determining which additives are necessary for a robust solar panel coating.

Finally, the power output data from the solar panel tests will determine how each wax formulation affects light transmission and energy efficiency. Higher voltage readings indicate that the wax blend allows for more light to reach the panel's surface, thus maintaining or even enhancing energy output. If a formulation with added UV stabilizers maintains higher power output while also preserving the physical structure, it may be ideal for applications where durability and efficiency are required. Conversely, if certain formulations result in lower power output, they may be too opaque, causing interference with the sunlight absorption necessary for energy generation.

By combining the results from melting point, density, aging, and power output data, the optimal wax formulation that balances thermal stability, UV resistance, and minimal interference with light transmission can be identified. For instance, if a blend with a higher ratio of carnauba wax and antioxidants shows high melting points, low degradation, and efficient solar intake, this formulation would likely be well-suited for protective applications on solar panels. Additionally, examining correlations between physical properties, such as melting point and thermal stability, and functional outcomes like power output, will further help to determine how these factors impact overall performance. For example, a higher melting point might correlate with better endurance under sunlight, suggesting that thermal stability is key to withstanding environmental conditions on solar panels. Likewise, density could affect light transmission, where a denser, more uniform coating might either enhance or reduce solar intake efficiency. Identifying these relationships will reveal the interplay between a formulation's physical characteristics and its suitability for outdoor applications, allowing you to make data-backed predictions about its long-term functionality. This thorough approach will help pinpoint a formulation that maximizes both the lifespan and efficiency of solar panels while enduring real-world environmental stress.

To ensure the reliability of the results, the data will be analyzed for consistency and accuracy.

Below are photographs of each wax and wax blend tested. From left to right: Pure Beeswax, Beeswax with Carnauba Wax in a 5:1 Ratio, Beeswax with Carnauba Wax 4:1 Ratio, Beeswax with Carnauba Wax in a 3:1 Ratio, Beeswax with UV Stabilizer, and Beeswax with Vitamin Power



Data Analysis and Results

Below is a table showing the melting points found of the wax blends containing only beeswax and carnauba wax. Beeswax has a melting point of about 145 degrees fahrenheit while carnauba wax has a melting point of 180 degrees fahrenheit (*Properties of Waxes* 4).

Wax Blend	Discovered Melting Point
Beeswax with Carnauba wax in a 5:1 Ratio	151
Beeswax with Carnauba wax in a 4:1 Ratio	160
Beeswax with Carnauba wax in a 3:1 Ratio	175

The data reflects the melting points of different mixtures of beeswax and carnauba wax at varying ratios. As the ratio of beeswax to carnauba wax decreases, meaning a higher proportion of carnauba wax is present, the melting point of the mixture increases. For instance, the melting point rises from 151°F at a 5:1 ratio to 160°F at a 4:1 ratio, and then to 175°F at a 3:1 ratio. This trend suggests a predictable relationship between the ratio of the two waxes and the thermal stability of the blend.

The observed increase in melting point can be attributed to the higher thermal stability of carnauba wax, which has a melting point of approximately 180°F compared to beeswax's lower melting point of around 145°F. This enhanced stability was evidenced not only by the elevated melting points but also by all the waxes' improved resistance to thermal degradation, such as discoloration, at higher temperatures. Carnauba wax, known for its high melting point and robust

structure, likely contributed to the stability of the blend by reinforcing its thermal resilience. In contrast, blends with a higher ratio of beeswax showed signs of thermal degradation, including a noticeable change in color when exposed to elevated temperatures, indicating less stability. These findings suggest that incorporating more carnauba wax can improve the thermal performance and durability of the wax mixture, making it a more suitable candidate for applications like solar panel coatings where consistent thermal properties are essential.

This steady increase in melting point and resistance to degradation indicates a potential linear relationship between the ratio of the waxes and their melting point, demonstrating that the higher concentration of carnauba wax a wax blend has, the better its thermal stability will be.

Beeswax with Vitamin E Powder:
• Color: Maintained original color, slower discoloration. • Texture: Stayed flexible but became bendy, did not prevent degradation. • Scent: Lost most scent. • Overall: Helped preserve flexibility, but did not stop degradation.
Beeswax with UV Stabilizer:
• Color: Retained original color, slower discoloration. • Texture: Became more flexible from solid to bendy. • Scent: Lost scent. • Overall: UV stabilizer preserved pliability and prevented hardening.
Pure Beeswax:
• Color: Darkened and yellowed over time. • Texture: Became more brittle. • Scent: Strong beeswax scent at first, then faded. • Overall: Degraded over time, lost flexibility and scent.

Accelerated aging tests simulate long-term environmental conditions in a short time to assess a material's durability and stability. Below are the observed results from the accelerated aging tests conducted on pure beeswax, beeswax blended with UV stabilizer, beeswax blended with tocopherol.

The results of the accelerated aging tests reveal differences in how Vitamin E powder, UV stabilizers, and pure beeswax perform under stress conditions. In terms of color retention, the beeswax with Vitamin E powder and UV stabilizers outperformed pure beeswax, which darkened and yellowed significantly over time. Beeswax with Vitamin E showed slower discoloration, maintaining a closer resemblance to its original appearance, while the UV stabilizer preserved the original color even more effectively, demonstrating its superior ability to counteract UV-induced changes.

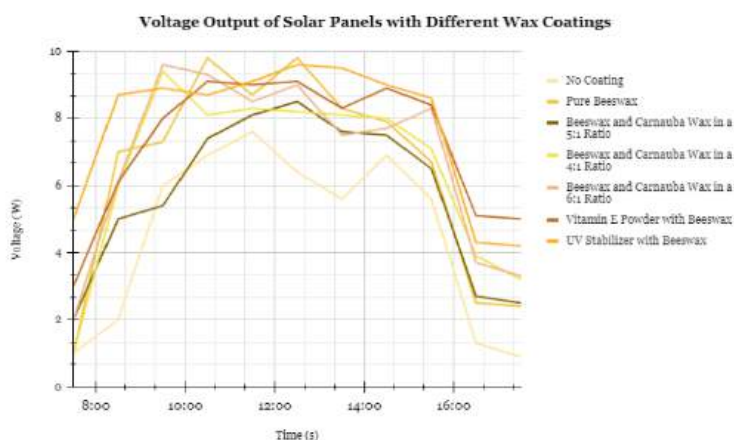
Texture and flexibility results further emphasize these differences. Pure beeswax became brittle as it aged, losing its flexibility entirely. In contrast, the beeswax treated with Vitamin E powder retained some flexibility but transitioned to a bendy state, indicating partial degradation. The UV stabilizer, however, stood out by preventing the wax from hardening; it became more pliable, transitioning from a solid state to a flexible, bendy texture, which suggests that UV stabilizers actively counteract the structural breakdown caused by prolonged UV exposure.

Scent preservation posed a challenge for all samples. Pure beeswax initially retained a strong, characteristic scent but lost it as degradation progressed. Similarly, both the Vitamin E and UV stabilizer-treated beeswax experienced significant scent loss over time, showing that

neither additive effectively protected against olfactory degradation. This uniform loss of scent suggests that additional preservation methods are required to maintain the sensory characteristics of beeswax.

Based on these findings, it was found that UV stabilizers are recommended for applications requiring enhanced flexibility and resistance to hardening under aging conditions. The UV stabilizer's ability to maintain flexibility while slowing discoloration makes it particularly suitable for situations where mechanical and aesthetic integrity are critical. However, the Vitamin E powder's contribution to slower discoloration indicates potential benefits from a combined approach. By blending UV stabilizers with Vitamin E powder, a synergistic effect could be achieved, addressing both flexibility and color preservation more effectively than either additive alone.

After the accelerated aging tests each wax blend was applied onto solar panels and measured for their voltage output. Below is a graph showing the results from the test.



The data indicates that solar panels coated with wax blends, particularly those incorporating Vitamin E powder or UV stabilizers, consistently achieved higher voltage output compared to panels coated with pure beeswax or left uncoated. This improvement can be attributed to the unique properties of the additives. Vitamin E powder, known for its antioxidant properties, likely mitigated the oxidative degradation of the wax coating caused by prolonged UV exposure. By maintaining the wax's structural and optical properties over time, Vitamin E allowed the solar panels to retain a smoother and more reflective surface, facilitating better light absorption and reducing energy losses.

Similarly, the panels with UV stabilizer-enhanced coatings demonstrated superior performance, showing the highest voltage readings throughout the day. UV stabilizers are specifically designed to absorb and dissipate harmful ultraviolet radiation, preventing structural damage to the wax and minimizing surface deformation or cracking. This preservation of surface integrity reduces light scattering, allowing more sunlight to penetrate the panel's active layers. Additionally, UV stabilizers may have contributed to thermal stability, preventing the wax from

softening or melting during peak sunlight hours, which could otherwise create uneven surfaces that diminish efficiency.

The beeswax and carnauba wax blends also improved voltage output compared to pure beeswax, with higher ratios of carnauba wax yielding better results, particularly in the afternoon. Carnauba wax has a higher melting point and better reflective properties than beeswax, making it more resilient to heat-induced softening and UV damage during periods of intense sunlight. These properties help maintain a stable, smooth coating, ensuring optimal light transmission to the solar panel. The slightly better performance observed in the afternoon suggests that carnauba wax plays a critical role in maintaining efficiency during the day's hottest hours when thermal stress is most pronounced.

In contrast, pure beeswax, while offering some level of protection, showed significant limitations. Beeswax tends to degrade under prolonged UV exposure and heat, leading to surface cracking, discoloration, and a reduction in optical clarity. This degradation likely caused increased light scattering and reduced the panel's ability to absorb sunlight effectively, resulting in lower voltage outputs over time.

The superior performance of the Vitamin E powder and UV stabilizer coatings can be explained by their ability to address the primary degradation pathways for wax coatings: UV radiation, oxidation, and thermal stress. By preserving the wax's optical clarity and structural integrity, these additives allowed the solar panels to function more efficiently. The beeswax and carnauba wax blends performed better than pure beeswax due to carnauba wax's inherent properties, such as its high melting point and durability, but they were less effective than the stabilizer-enhanced coatings because they lacked active components to counteract UV and oxidative damage.

To maximize solar panel efficiency, wax coatings should incorporate stabilizing additives like Vitamin E powder or UV stabilizers. These additives enhance the durability of the wax, maintaining surface smoothness and reflectivity under prolonged sunlight exposure. Further testing could explore optimizing the ratios of beeswax and carnauba wax in conjunction with stabilizers to balance cost and performance.

Discussion and Conclusion

This study has provided an exploration of the role of wax coatings in enhancing the performance and longevity of solar panels. By evaluating the effects of melting point, thermal stability, and solar panel output, the findings reveal how material composition and additives interact to influence the overall effectiveness of wax coatings in solar energy applications.

The melting point data demonstrates the importance of wax composition in determining how well the coating can withstand heat. As the proportion of carnauba wax in the blend increased, the melting point also rose, indicating that carnauba's higher melting point provides enhanced thermal stability. This is particularly significant for solar panels, which are subjected to fluctuating temperatures throughout the day. The ability of the wax to retain its structural

integrity under high heat prevents the coating from softening or distorting, thus maintaining a smooth and reflective surface essential for efficient light absorption.

When considering the accelerated aging results, the wax blends with Vitamin E powder and UV stabilizers performed markedly better than pure beeswax. Vitamin E, by slowing oxidative degradation, and UV stabilizers, by preventing UV-induced breakdown, collectively helped the coatings maintain their flexibility and thermal resilience over time. Pure beeswax, in contrast, showed more rapid signs of degradation, including increased brittleness and discoloration. This highlights the critical need for stabilizing additives in enhancing the thermal stability and longevity of wax coatings.

The results from the solar panel tests indicate that wax coatings significantly affect the overall efficiency of solar panels, with the highest voltage outputs consistently observed in panels coated with Vitamin E powder and UV stabilizers. These coatings helped preserve the optical properties of the wax, reducing light scattering and maximizing sunlight absorption. The more stable the wax coating is, the more sunlight the panel can effectively capture and convert into electricity. This is a clear advantage over panels with no coating or those coated with pure beeswax, which suffer from degradation over time, resulting in reduced voltage outputs.

Interestingly, the beeswax and carnauba wax blends also enhanced performance compared to pure beeswax, especially during the afternoon when sunlight is at its peak. The higher the proportion of carnauba wax, the better the afternoon performance, likely due to carnauba's superior reflectivity and resistance to heat. This reveals that the intrinsic properties of carnauba wax, such as its high melting point and structural rigidity, can be further optimized through synergistic interactions with stabilizing agents. However, even with these improvements, the performance of these blends was not as consistent or high as those with stabilizing additives, underscoring the importance of incorporating additives to optimize long-term solar panel efficiency.

Taken together, the findings from the melting point, thermal stability, and solar panel output tests reveal the intricate relationship between material composition, additives, and solar panel performance. The combination of waxes and stabilizers provides a multifaceted solution to the challenges posed by prolonged exposure to sunlight. While beeswax improves thermal resilience and reflectivity, it is the additives—Vitamin E powder and UV stabilizers—that significantly enhance the wax's durability and effectiveness in preserving both optical properties and structural integrity. These factors work synergistically to ensure that the solar panels can operate efficiently throughout the day, with minimal loss in performance over time.

Furthermore, the importance of thermal stability and the preservation of optical properties cannot be overstated. Solar panels exposed to high heat need coatings that can maintain their structure and reflective capabilities, preventing them from becoming brittle or distorted. Additives that provide UV protection and prevent oxidation are crucial in achieving this goal, as they actively counter the degradation that would otherwise reduce the lifespan and efficiency of the wax coating. These additives enhance the wax's molecular stability by neutralizing free

radicals and reducing photodegradation, thereby prolonging its functional lifespan under UV exposure and high-temperature conditions.

This suggests that the composition of wax blends is not merely a function of their base materials but also a dynamic system influenced by external modifiers. In conclusion, the study reinforces the idea that wax coatings, especially those incorporating stabilizing additives like Vitamin E powder and UV stabilizers, offer significant benefits for enhancing the efficiency and durability of solar panels. These coatings not only improve voltage output by preserving the optical and structural properties of the wax but also contribute to the long-term stability of the solar panels, reducing the effects of thermal and UV-induced degradation. While beeswax and carnauba wax blends provide a notable improvement over pure beeswax, the most effective performance was observed in coatings enhanced with stabilizers.

The findings point to the importance of a holistic approach in the design of solar panel coatings—one that integrates both the right material composition and stabilizing additives to address the complex challenges of UV exposure, thermal stress, and long-term durability. Future research should focus on optimizing these wax blends for a balance between cost and performance, exploring different ratios of waxes and additives, and testing under more diverse environmental conditions. Ultimately, the integration of these advanced coatings could lead to more efficient, durable, and sustainable solar panels, contributing to the broader goal of renewable energy advancement.

Works Cited

- Aljaghoub, H., Abumadi, F., AlMallahi, M., Obaideen, K., & Alami, A. H. (n.d.). Solar PV cleaning techniques contribute to Sustainable Development Goals (SDGs) using Multi-criteria decision-making (MCDM). *International Journal of Thermofluids*, 16. <https://doi.org/10.1016/j.ijft.2022.100233>.
- Aydin, A., & Kocak, S. (2021). Thermal stability and chemical properties of natural wax blends for energy applications. *Journal of Applied Polymer Science*, 138(15), 50040. <https://doi.org/10.1002/app.50040>
- Better Health Channel. (n.d.). Antioxidants. *Better Health Channel*. <https://www.betterhealth.vic.gov.au/health/healthyliving/antioxidants>
- Creative Proteomics. (n.d.). Waxes: Properties, functions, and analysis techniques. *Creative Proteomics*. <https://lipidomics.creative-proteomics.com/resource/waxes-properties-functions-analysis.htm>
- Dincer, I., & Rosen, M. A. (2016). *Phase Change Materials: Science and Applications*. New York, NY: Springer.
- Fath, H. E., & Dakhil, H. M. (2019). A review of solar energy applications and their impact on energy savings. *Renewable and Sustainable Energy Reviews*, 101, 1-17. <https://doi.org/10.1016/j.rser.2018.09.028>
- Foreverest Resources Ltd. (n.d.). The natural attributes of waxes. *Foreverest Resources Ltd*. <https://foreverest.net/news-list/the-natural-attributes-of-waxes>
- Futahashi, R., Yamahama, Y., Kawaguchi, M., Mori, N., Ishii, D., Okude, G., Hirai, Y., Kawahara-Miki, R., Yoshitake, K., Yajima, S., Hariyama, T., & Fukatsu, T. (2019). Molecular basis of wax-based color change and UV reflection in dragonflies. *eLife*, 8, e43045. <https://doi.org/>
- Srimanickam, B., Raja, N. D., Saranya, A., & others. (2024). Evaluating the effectiveness of paraffin wax and zinc sulphide nanoparticles to improve the performance of hybrid solar panel/dryer. *Interactions*, 245(201). <https://doi.org/10.1007/s10751-024-02013-w>
- International Renewable Energy Agency. (2020). *Renewable Power Generation Costs in 2020*. <https://www.irena.org/publications/2021/Jun/Renewable-Power-Costs-in-2020>
- Isimjan, T. T., & West, D. H. (2023). Do functional coatings work for solar panel dust mitigation? *ACS Materials Letters*, 5(10), 2718–2725. <https://doi.org/10.1021/acsmaterialslett.3c00619>
- Jampole, E., & Chaudhary, S. (n.d.). *Shedding Light on Solar Panel Degradation*. Exponent. Retrieved December 26, 2024, from <https://www.exponent.com/article/shedding-light-solar-panel-degradation>
- Lackner, K. S., & Sachs, J. D. (2005). A robust strategy for sustainable energy. *Brookings Papers on Economic Activity*, 2005(2), 215-269.

- Law, A. M., Jones, L. O., & Walls, J. M. (2023). The performance and durability of anti-reflection coatings for solar module cover glass – A review. *Solar Energy*, 261, 85–95. <https://doi.org/10.1016/j.solener.2023.06.009>
- Masterson, V. (2022, August 9). *Why don't solar panels work as well in heatwaves?* World Economic Forum. Retrieved December 26, 2024, from <https://www.weforum.org/stories/2022/08/heatwaves-can-hamper-solar-panels/>
- Mozumder, M. S., Mourad, A.-H. I., Pervez, H., & Surkatti, R. (2019). Recent developments in multifunctional coatings for solar panel applications: A review. *Solar Energy Materials and Solar Cells*, 189, 75–102. <https://doi.org/10.1016/j.solmat.2018.09.015>
- NIOSH Pocket Guide to Chemical Hazards. (n.d.). *Waxes*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. <https://www.cdc.gov/niosh/npgd/npgd0197.html>
- Occupational Safety and Health Administration. (n.d.). *Safety and Health Topics: Chemical Safety*. <https://www.osha.gov/chemical>
- Perrakis, G., Tasolamprou, A. C., Kenanakis, G., Economou, E. N., Tzortzakis, S., & Kafesaki, M. (2020). Ultraviolet radiation impact on the efficiency of commercial crystalline silicon-based photovoltaics: A theoretical thermal-electrical study in realistic device architectures. *OSA Continuum*, 3(6), 1436–1444. <https://doi.org/10.1364/OSAC.3.001436>
- Properties of Waxes*. (n.d.). Wax Emulsions. Retrieved March 21, 2025, from <https://www.wax-emulsions.com/properties-of-waxes/>
- Swartz, N., & Clare, T. L. (2015). ON THE PROTECTIVE NATURE OF WAX COATINGS FOR CULTURALLY SIGNIFICANT OUTDOOR METALWORKS. *Journal of the American Institute for Conservation*, 54(3), 181-201.
- TeKrony, D. M. (2005). Accelerated aging test: Principles and procedures. *Seed Technology*, 27(1), 135-146.
- Teusink, R. S., Rahman, M., Bressan, R. A., & Jenks, M. A. (2002). Cuticular waxes on *Arabidopsis thaliana* close relatives *Thellungiella halophila* and *Thellungiella parvula*. *International Journal of Plant Sciences*, 163(2), 309-315. <https://doi.org/10.1086/338825>
- Turner, L. (2017). A Roof over Your Head. *ReNew: Technology for a Sustainable Future*, (138), 52-60. JSTOR. <https://www.jstor.org/stable/renetechsustfutu.138.52>
- United States Environmental Protection Agency. (2021). *Managing Hazardous Waste*. <https://www.epa.gov/hw/managing-hazardous-waste>
- Wax in Biochemistry: Properties, Functions, and Contrasts with Triglycerides*. (n.d.). Creative Proteomics. Retrieved December 27, 2024, from <https://www.creative-proteomics.com/resource/waxes-biochemistry-functions-contrasts.htm>
- Xie, Y., Sengupta, M., Habte, A., & Andreas, A. (2022). The "Fresnel equations" for diffuse radiation on inclined photovoltaic surfaces (FEDIS). *Renewable and Sustainable Energy Reviews*, 161, 112362. <https://doi.org/10.1016/j.rser.2022.112362>

Mental Health Trends in Just Dance By Manya Jain

Abstract

The dancing industry has been studied to a heavy extent with understanding mental health patterns, within the professional industry. Something lacking is studying the mental health impact of dancing as a hobby. This study examines the community of Just Dance and its emotional impact. Especially when it comes to professional dancing and mental health, much research depicts some dancers are languishing or some are thriving and flourishing (1). This study's purpose is to look into different aspects of dancing, mainly an activity that people learn using platforms like Just Dance, to see if its impact on mental health is positive or negative.

Introduction

As many young adults and even kids have grown up and relied on technology for a while, Just Dance has been an enjoyable platform where many can increase their physical activity. It is important to note the benefits that can come from consistent dancing, improved cardiovascular health, balance, strength, flexibility, spatial awareness, etc [2]. Many research shows that as much as dance improves and challenges one's body, it also improves the brain.

Some areas the brain focuses on amid dancing are fast decision-making, focused attention, maintaining immediate balance, rapid problem-solving, coordination, and much more [2]. Areas like focused attention and rapid problem-solving are significant areas that Just Dance helps. With the system of earning points, the more one's movement aligns with the dancer on screen, the more heavily it contributes to brain tasks of attention and problem-solving. In our study, we explore the emotions users feel with the game Just Dance. Emotions contribute significantly to the status of mental health, and this can help us understand the psychological effects of movement-based games on players. Positive emotions, or engaged players may indicate benefits such as stress relief and increased dopamine levels, while negative emotions, or disengaged players may indicate parallels with stress, challenges with the game, etc. Most of the brain is active and multitasking while following complex dance movements [2].

Methodology

In our study, we use a Kaggle dataset that documents the commentary of Just Dance Players. With these comments, we run sentiment analysis and BERT models to study the significance of the resulting emotions.

Results

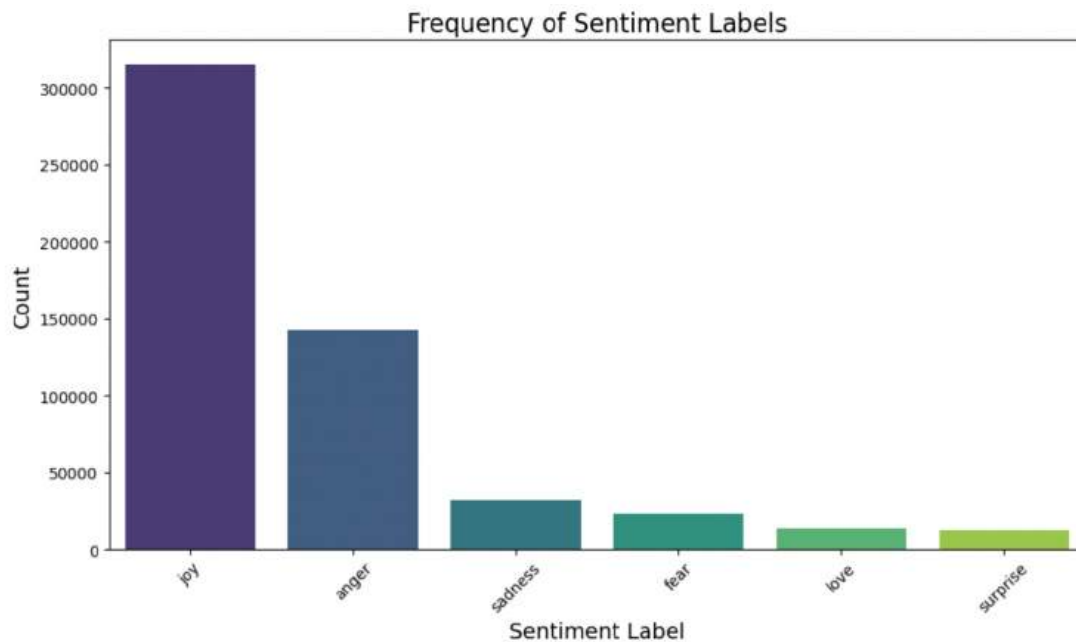


Figure 1

In this resulting bar graph of emotions, there is a high number of the emotion “joy.” Something we noted was how anger was an emotion that came second. Anger is one of the most challenging emotions to address in specific areas, and dance can mean a wide range of things [3]. Anger commonly results from external factors that the person believes they shouldn’t have experienced [3]. With Just Dance, anger could result from mistakes in following through with the supposed dance move, competition with friends, the point system, and more. This anger could mean something harmless that can motivate players to participate more and get mental revenge on the external event that originated their anger.

With the emotions of joy, anger, sadness, etc, the emotional chemicals of dopamine and norepinephrine are released which enhance long-term learning [2]. Platforms like Just Dance help promote such neurological improvements. Studies have previously shown that participation in dance programs can have a positive impact on all ages [4]. Dancing proves to show that people come out of their comfort zone, improve their bodily health, and more. With the high level of satisfaction shown in the emotions, Just Dance continues to be a platform that can significantly benefit one’s mental and physical health.



Figure 2

In our study, we conducted a word cloud to analyze whether or not the language players message each other can affect the emotional outcome of playing Just Dance. The number one word that was most popular resulted in being positive. When players exhibited their feelings and messages post-dancing, it seemed that positive was the word that appeared the most. The majority of players in the study seemed pleasantly satisfied with the game. More of the common words were “love, beautiful, good, fire, and like” most of which are positively associated with emotional health.

Conclusion

Overall, there is a common thread of users enjoying the world of Just Dance. From our results, the majority of players have been positive and promoted dancing. Just Dance is a platform of enjoyment and beneficial for mental health. Our findings highlight the intense emotional impact of Just Dance on the players, with joy and engagement dominating the sentiment analysis. There was frustration and anger in some cases, but these typically emerged from gameplay difficulty and not from negative emotional results. The overall positive reactions, as reflected in the word cloud analysis, suggest that Just Dance provides a fun and mentally stimulating environment.

By incorporating gaming with exercise, Just Dance enhances cognitive functions such as problem-solving concentration, and emotional well-being. That positive emotional responses are forthcoming, together with the neurological benefits of movement, suggests that dance—either as a vocation or an avocation—can be important in assisting mental health. Future studies can continue to explore how different gaming settings impact emotional outcomes and whether games like Just Dance can be integrated into mental health interventions.

Works Cited

- Millman, LS Merritt, et al. "Towards a neurocognitive approach to dance movement therapy for mental health: A systematic review." *Clinical Psychology & Psychotherapy* 28.1 (2021): 24-38. [4]
- Parise, Maria. "The intersection of anger management counseling and embodiment in dance/movement therapy." (2017). [3]
- Salo, Amber, "The Power of Dance: How Dance Effects Mental and Emotional Health and Self-Confidence in Young Adults" (2019). Master's Theses. 133.
<https://digscholarship.unco.edu/theses/133> [2]
- Schachtler, Dwarika Michelle, and Haraldsen Heidi Marian. "Mental Health in Dance: A Scoping Review." *Frontiers in Psychology*, vol. 14, 2023,
<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1090645>.
Accessed 6 Jan. 2025. [5]
- Sivvas, Georgios, et al. "Dance contribution in health promotion." *Journal of physical education and sport* 15.3 (2015): 484. [1]

AI-Powered Tax Insights and Predictions For a Multilingual Audience

By Malcolm L. Gao and Chenge Zhang

Abstract

Economic disparities of non-English speaking audiences and particularly of the immigrants with a diverse and multilingual background can exacerbate dramatically when it comes to the tax reporting season. The language barrier has been identified as one leading reason that prevents countless US households from receiving their due tax benefits and justice. In recent years, significant breakthroughs in artificial intelligence (AI) and machine learning (ML) are expected to transform various aspects of daily lives. In particular, large language models (LLMs) have been utilized across numerous domains. In this study, we leverage AI models to assist the general public, including individuals with diverse language backgrounds, in understanding the latest news and policies from the Internal Revenue Service (IRS). Specifically, we employ the Llama-2-7b-chat-hf model, an open-source large language model developed by Meta (<https://llama.meta.com>), to automatically generate summaries and Q&A (question-and-answer) pairs from tax-related news. Additionally, we use the AI model to detect and verify potential fake tax news. We also explore the use of a machine learning model, Scikit-learn, to predict future tax trends based on previous years' data. To implement these functionalities, we have developed a set of Python scripts in a Linux Red Hat 9.4 environment. Furthermore, our Python code includes a feature that converts text output into MP3 audio versions, catering to users who prefer auditory information over reading materials. The protocol and usage guidelines have been documented on GitHub for public access (<https://github.com/TaoTao70580/AI-ToolBox>). Limitations and future suggestions will also be discussed.

Introduction

The US has witnessed a rapid demographic change in the past decade, and by 2019 about 20% of the households speak a foreign language (Zeigler and Camarota, 2019). Economic disparity among all people and particularly within immigrants grows rapidly (Zan et al., 2024), and out of many sociological and political reasons, the language barrier has been found to be one of the leading factors responsible for earning differences (Chiswick and Miller, 1999). Particularly, it is noted that when it comes to taxes, the general public faces challenges in understanding complex tax rules and regulations and often misses the latest tax changes. This issue is even more pronounced for non-English-speaking taxpayers, who struggle to keep up with tax updates in the news. Additionally, the public is frequently confused by inconsistencies in media reports and even false or misleading tax information found online. A key reason for this susceptibility to misinformation is the language barrier between professional tax articles and daily language, as well as the difficulty in distinguishing between mainstream interpretations and isolated, ad hoc cases. It is clearly a daunting challenge to mitigate such knowledge gap and ultimately the resulting income disparity. In one recent study, it is found that a school-based outreach to students and their non-English speaking parents dramatically improves the

knowledge on taking the child tax credit (Łuczywek et al., 2024). Information page provided by IRS is rather not flexible for individual needs (IRS Tax Tip 2023-43). Admittedly, the depth and breadth of such outreach activities are very limited, and more efficient approaches are awaited that can make up the knowledge gap through simple tools.

Recently, we have witnessed significant advancements in Artificial Intelligence (AI) and Machine Learning (ML), driven by research institutions and major tech hubs like Silicon Valley and beyond. Motivated by these challenges and technological progress, this study aims to develop an AI-powered toolbox capable of analyzing and summarizing complex IRS tax rules and regulations while generating summary and question-and-answer (Q&A) pairs for public understanding. To achieve this, we leverage large language models (LLMs), which have demonstrated success in various applications. Specifically, we employ Meta’s Llama-2-7b-chat-hf model (<https://huggingface.co/meta-llama/Llama-2-7b-chat-hf>) and have developed a set of Python scripts to analyze tax-related articles. Our model not only translates tax summaries and Q&As into multiple languages but also processes tax articles written in different languages. Furthermore, taking advantage of the first author’s proficiency in English and Spanish and the second author’s expertise in English and Chinese, we have tested our model in English, Spanish, and Chinese—three of the world's most widely spoken languages. Additionally, we employ AI to detect and verify potential fake tax news. We also explore the use of the Scikit-learn machine learning model to predict future tax trends based on previous years' data. Finally, we discuss existing gaps in our AI toolbox and highlight future research directions and improvements.

Methods

Our Python codes for the proposed AI workflow were developed in the Linux Red Hat 9.4 in a virtual python environment. To enable the python environment, one needs to run the following commands.

```
python3 -m venv venv  
source venv/bin/activate
```

Fig. 1: Command window for the python environment.

Next, one needs to install required python libraries “transformers” and “torch”, as can be found from <https://pypi.org/project/transformers/> and [https://pypi.org/project/touch/](https://pypi.org/project/torch/). Command to install required libraries and other packages are shown below in Fig. 2.

```
pip install transformers torch
```

```
pip install huggingface_hub  
pip install sentencepiece  
pip install gtts  
pip install scikit-learn  
pip install pandas
```

Fig. 2: Common window for the installation of python libraries.

The employed Llama-2-7b-chat-hf model typically works well with natural languages for summary and Q&A generation. However, tax publications often include tables. Tables are effective for human readers to understand structural content but less suitable for large language models to process. To address this challenge, the authors developed a python code to call OpenAI API (gpt-3.5-turbo-instruct) to transform tables into descriptive paragraphs. The complete code can be found under authors's Github page:

<https://github.com/TaoTao70580/AI-ToolBox/blob/main/parse.py>

Next, the code will load the Llama-2-7b-chat-hf model and tokenizer. Read the article and generate a summary. The similar method can be used to generate Q&A pairs. The complete code can be found under authors's Github page: <https://github.com/TaoTao70580/AI-ToolBox>

```
# Load the tokenizer and model  
model_name = "meta-llama/Llama-2-7b-chat-hf"  
tokenizer = AutoTokenizer.from_pretrained(model_name)  
model = AutoModelForCausalLM.from_pretrained(model_name)  
# Ensure the model is on the right device  
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")  
model.to(device)  
  
# Function to generate summary  
def generate_summary(text, tokenizer, model, device):  
    prompt = f"Generate a summary in two or three sentences based on the  
following tax article:\n\n{text}\n\n "  
    inputs = tokenizer(prompt, return_tensors='pt').to(device)  
  
    with torch.no_grad():  
        outputs = model.generate(inputs['input_ids'], max_length=600,  
num_return_sequences=1, temperature=0.7, num_beams=5)  
  
    generated_text = tokenizer.decode(outputs[0], skip_special_tokens=True)  
    return generated_text
```

Fig. 3: Procedure to load the model and tokenizer.

The next step is to translate output into other languages

(<https://github.com/TaoTao70580/AI-ToolBox/blob/main/translate.py>).

```
# Load the Llama model for reading
llama_model_name = "meta-llama/Llama-2-7b-chat-hf"
llama_tokenizer = AutoTokenizer.from_pretrained(llama_model_name)
llama_model = AutoModelForCausalLM.from_pretrained(llama_model_name)

# Load translation models
translation_model_es =
MarianMTModel.from_pretrained('Helsinki-NLP/opus-mt-en-es')
translation_tokenizer_es =
MarianTokenizer.from_pretrained('Helsinki-NLP/opus-mt-en-es')
translation_model_zh =
MarianMTModel.from_pretrained('Helsinki-NLP/opus-mt-en-zh')
translation_tokenizer_zh =
MarianTokenizer.from_pretrained('Helsinki-NLP/opus-mt-en-zh')

# Ensure the models are on the correct device
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
llama_model.to(device)
translation_model_es.to(device)
translation_model_zh.to(device)

# Function to translate text
def translate_text(text, tokenizer, model, device):
    inputs = tokenizer(text, return_tensors="pt", padding=True,
truncation=True).to(device)
    translated_tokens = model.generate(**inputs)
    translated_text = tokenizer.batch_decode(translated_tokens,
skip_special_tokens=True)[0]
    return translated_text
```

Fig. 4: Output step.

The following code (<https://github.com/TaoTao70580/AI-ToolBox/blob/main/voice.py>) will convert text output into MP3 for users who prefer auditory information.

```
from gtts import gTTS
import os
mytext = "Your_input_txt "

language = 'en'
myobj = gTTS(text=mytext, lang=language, slow=False)

myobj.save("output.mp3")
```

Fig. 5: The auxiliary step for multimedia interactions.

The AI model in the following code will detect misinformation in the input. The complete code can be found at: https://github.com/TaoTao70580/AI-ToolBox/blob/main/check_fake_info.py

```
# Function to detect misinformation
def assess_statement(text, tokenizer, model, device):
    prompt = (
        f"Evaluate the following statement for potential misinformation or falsehood. "
        f"Consider known facts and current understanding in your assessment. "
        f"Provide a brief explanation if the statement is misleading or false.\n\n"
        f"Statement: {text}\n\n"
        f"Assessment:"
    )
    inputs = tokenizer(prompt, return_tensors='pt').to(device)

    with torch.no_grad():
        outputs = model.generate(inputs['input_ids'], max_length=500, num_return_sequences=1,
                                temperature=0.7, num_beams=5)

    generated_text = tokenizer.decode(outputs[0], skip_special_tokens=True)
    return generated_text
```

Fig. 6: The step of misinformation examination.

Machine learning models which can be integrated into Python scripts could be trained to predict future tax liabilities based on historical data. This can help the user plan ahead for the next tax year by forecasting changes in tax liabilities. The following example uses machine learning (Linear Regression) to predict tax owed based on previous income and deductions. The authors hope that college students, tax professionals, and social volunteers will have interests in this task. With the needed access to computing resources with sufficient GPUs to run AI training jobs, it is anticipated that the model will be trained on historical tax data, used to predict future income tax liabilities based on various factors (e.g., changes in income, deductions, or tax laws), ultimately beneficial to the multilingual audience.

```
# Simple machine learning model to predict future tax based on previous year
from sklearn.linear_model import LinearRegression

def predict_future_tax(previous_year_income, previous_year_deductions):
    # Assume we have historical data on income and deductions
    X = [[50000, 12000], [60000, 14000], [70000, 15000]] # Example data: [income, deductions]
    y = [8000, 8500, 10000] # Example tax owed for each case

    model = LinearRegression()
    model.fit(X, y)

    # Predict tax for the next year based on current income and deductions
    prediction = model.predict([[previous_year_income, previous_year_deductions]])
    return prediction[0]

predicted_tax = predict_future_tax(65000, 14500)
print(f"Predicted tax for next year: ${predicted_tax:.2f}")
```

Fig. 7: Machine learning module with tax prediction as an example.

By combining the above method with the following Python code to calculate tax, one can see that AI can assist the general public with tax-related tasks.

```
import pandas as pd

# Example tax data (income, deductions, tax withheld)
tax_data = {
    'Total_Income': 60000,
    'Deductions': 14200,
    'Tax_Withheld': 8000
}

# Step 1: Calculate taxable income
taxable_income = tax_data['Total_Income'] - tax_data['Deductions']

# Step 2: Define tax brackets (single filer)
brackets = [(0, 11000, 0.10), (11000, 44725, 0.12), (44725, 95375, 0.22)]

# Step 3: Calculate tax owed based on brackets
def calculate_tax(taxable_income, brackets):
    tax_owed = 0
    previous_bracket_upper = 0

    for lower, upper, rate in brackets:
        if taxable_income > upper:
            tax_owed += (upper - lower) * rate
        else:
            tax_owed += (taxable_income - previous_bracket_upper) * rate
            break
        previous_bracket_upper = upper
    return tax_owed

# Calculate tax owed
tax_owed = calculate_tax(taxable_income, brackets)

# Step 4: Calculate tax refund or due
tax_refund_or_due = tax_data['Tax_Withheld'] - tax_owed

#tax_owed, tax_refund_or_due
print ("Predicted tax owned : ", tax_owed)
print ( "Predicted tax refund or due : ", tax_refund_or_due)
```

Fig. 8: Demonstration of tax assistance.

The complete code set can be found under Github developed by the authors, including:
<https://github.com/TaoTao70580/AI-ToolBox/blob/main/parse.py>
https://github.com/TaoTao70580/AI-ToolBox/blob/main/q_and_a.py
<https://github.com/TaoTao70580/AI-ToolBox/blob/main/summary.py>
<https://github.com/TaoTao70580/AI-ToolBox/blob/main/translate.py>
<https://github.com/TaoTao70580/AI-ToolBox/blob/main/voice.py>

https://github.com/TaoTao70580/AI-ToolBox/blob/main/check_fake_info.py
https://github.com/TaoTao70580/AI-ToolBox/blob/main/cauculate_tax.py
https://github.com/TaoTao70580/AI-ToolBox/blob/main/future_tax.py

Results

Example 1: We have tested our Python codes with the following tax article

<https://www.wsj.com/personal-finance/taxes/1099-form-income-taxes-side-hustles-14e82f9f>

Those who never report their online income may face steep penalties. The IRS has launched an investigation into taxpayers who earned money as experts on the online platform JustAnswer, and allegedly failed to accurately report income from 2017 through 2020. At that time JustAnswer experts such as veterinarians and mechanics were generally paid between \$15 and \$25 for each question they responded to.

A judge in December authorized the IRS to issue a summons requiring JustAnswer to provide the names of anyone who earned \$5,000 or more on the platform in any one year in that time period, when the company apparently wasn't sending out 1099 forms. The company didn't respond to requests for comment.

Q: What is the purpose of the IRS investigation?

A: The IRS is investigating taxpayers who earned money as experts on the online platform JustAnswer and allegedly failed to accurately report income from 2017 through 2020.

Q: How much did JustAnswer experts typically earn per question?

A: According to the article, JustAnswer experts such as veterinarians and mechanics were generally paid between \$15 and \$25 for each question they responded to.

Q: What did a judge authorize the IRS to do in December?

A: A judge authorized the IRS to issue a summons requiring JustAnswer to provide the names of anyone who earned \$5,000 or more on the platform in any one year in that time period, when the company apparently wasn't sending out 1099 forms.

Fig. 9: One example and the code that generates the Q/A pairs.

In the next example, the code translates the above tax article into Chinese and Spanish:

Translated text in Chinese:

那些从未报告其在线收入的人可能会面临巨额罚款。国税局已经对在网上平台“Just Answer”上挣钱的纳税人进行了调查,并据称没有准确报告2017年至2020年的收入。当时,只有兽医和机械师等“安斯沃”专家在回答每个问题时通常得到15美元至25美元的工资。12月,一位法官授权国税局发出传票,要求Just Answer提供该平台在任何一年内挣5 000美元或更多美元的人的姓名,因为当时该公司显然没有发出1099张表格。该公司没有回应要求发表评论。

Translated text in Spanish:
 Los que nunca reportan sus ingresos en línea pueden enfrentar fuertes penalizaciones. El IRS ha iniciado una investigación sobre contribuyentes que ganaron dinero como expertos en la plataforma en línea JustAnswer, y supuestamente no reportó con precisión los ingresos desde 2017 hasta 2020. En ese momento, los expertos de JustAnswer, como veterinarios y mecánicos, fueron generalmente pagados entre \$15 y \$25 por cada pregunta a la que respondieron. Un juez en diciembre autorizó al IRS a emitir una citación que requería que JustAnswer proporcionara los nombres de cualquiera que ganara \$5,000 o más en la plataforma en cualquier año en ese período de tiempo, cuando la compañía aparentemente no estaba enviando 1099 formularios. La compañía no respondió a las solicitudes de comentarios.

Fig. 10: The customized output in various languages.

Example 2: We have tested our Python codes with the following fake tax news

The IRS has announced that all federal income tax will be suspended for the next five years, effective immediately. This new policy is expected to stimulate the economy by allowing taxpayers to keep more of their income.

Assessment:

The statement is false. There is no evidence to suggest that the IRS has announced a suspension of federal income tax, nor is there any indication of such a policy change on the IRS website or other reputable sources. In fact, the IRS has consistently maintained that federal income tax is a legally required payment and that failure to pay can result in penalties and fines. Therefore, this statement is misleading and should be disregarded.

Fig. 11: Another example demonstrating the detection of misinformation.

Example 3: Assume we have historical data on income and deductions. If the historical data on income and deductions are given in Fig. 12 (top), run the code in https://github.com/TaoTao70580/AI-ToolBox/blob/main/future_tax.py, which predicts the tax for next year in Fig. 12 (bottom).

[income, deductions]	[\$50000, \$12000]	[\$60000, \$14000]	[\$70000, \$15000]
Tax owed	\$8000	\$8500	\$10000

previous year income	previous year deductions	predicted tax for next year
\$65000	\$14500	\$9250
\$75000	\$14500	\$11750.00
\$85000	\$14500	\$14250.00

Fig. 12: The third example demonstrating the tax prediction.

Discussions

AI can significantly enhance the above Python code in various ways. For example, if a user provides tax information in plain language (e.g., 'I made \$60,000 last year and had \$14,200 in deductions'), an AI model like GPT-4 can parse the input and convert it into structured data for tax calculations. Additionally, GPT models can provide personalized tax advice. Once taxable income and deductions are determined, AI can suggest tax-saving strategies or explain applicable deductions based on the user's specific situation. For example, after calculating the tax, GPT can generate a personalized tax advice report.

However, this study is constrained by limited computational resources and relies solely on publicly available tax articles, which presents challenges in achieving sufficient training and iterative optimization—challenges we hope to address with future resources. While open-access IRS public documentation is available, initial estimates suggest that the required database and computing power exceed the capabilities of a single PC or laptop. Additionally, this study utilizes free AI model instead of the latest proprietary models, which may provide more accurate results.

Beyond model training and data processing, designing an interactive web-based user interface—such as a 'TaxGPT'-style chatbot—would enhance accessibility for users with tax-related questions. Technically, this would involve developing backend algorithms to process queries in real-time. With additional resources and collaboration, AI tools could be more effectively leveraged to improve tax services for people across diverse language backgrounds.

Works Cited

AI Toolsets as utilized by this study:

- <https://llama.meta.com/>
- <https://huggingface.co/TheBloke/Llama-2-7B-Chat-GGUF>
- <https://platform.openai.com/docs/models>

The Wall Street Journal. The IRS now knows a lot more about what you sell. Online on 1/21/2025.

<https://www.wsj.com/personal-finance/taxes/1099-form-income-taxes-side-hustles-14e82f9f>

Chiswick BR, Miller PW. Language skills and earnings among legalized aliens. *Journal of Population Economics*, 12: 63-89, 1999. <https://doi.org/10.1007/s001480050091>

Internal Revenue Services. IRS offers tax help in a variety of languages. IRS Tax Tip 2023-43, April 3, 2023.

<https://www.irs.gov/newsroom/irsgov-offers-tax-help-in-a-variety-of-languages>

Łuczywek B, Lippold K, Betts J, Berman E. Boosting take-up of the expanded child tax credit through school-based outreach. *National Tax Journal*, 77: 835-859, 2024.

<https://doi.org/10.1086/732681>

Zan H, Fan JX, Lozada B. The economic disparity between Hispanic and non-Hispanic white households: an analysis of middle-class achievement. *The American Journal of Economics and Sociology*, 83: 93-107, 2024. <https://doi.org/10.1111/ajes.12521>

Zeigler K, Camarota SA. 67.3 million in the United States spoke a foreign language at home in 2018. Center for Immigration Studies, October 2019.

https://cis.org/sites/default/files/2019-10/camarota-language-19_0.pdf

Assessing the Impact of the Hungry Healthy Free Kids Act on Child Health and Food Security By Justin Chen

Abstract

This study examines the effectiveness of the Hungry Healthy Free Kids Acts of 2010 in improving child health and food security outcomes, focusing on key indicators such as obesity rates, death rates, food security rates, and lunches served. The study uses the t-test and regression analysis to compare various indicators before and after policy implementation. The results show that the Hungry Healthy Free Kids Act was successful in reducing the probability of dying among the youth and was effective in ensuring participation in the program, even among lower-income students who required free or reduced-priced meals. However, the policy was not effective in reducing the percentage of overweight and obese youth in the United States or in reducing the percentage of households with food-insecure children.

Keywords

School Lunch, T-test, Children Health, HHFKA, Regression

Introduction

The Hungry Healthy Free Kids Acts (HHFKA), implemented in 2010, aimed to improve child health and food access by enhancing the national standard of school meals to align with the 2010 Dietary Guidelines for Americans, increase access to these lunches, and combat child obesity (Cohen et al., 2014).

The Healthy, Hunger-Free Kids Act of 2010 allowed for \$4.5 billion in funding over the next 10 years towards federal school meal and child nutrition programs. Its objective was to increase access to healthy food for low-income children, curb the rising rate of obesity in the United States, and ensure that school meals were closely inspected and aligned with nutritional guidelines. The Act serves to improve child health and ensure access to healthy food by making changes to school lunches. (White House, 2010⁸).

Much of the youth population consume school lunches (see Child Nutrition Data, 2025⁹), which affects their diet, health, and ultimately their prosperity. Because the policy will impact the well-being of current and future generations, it is critical to evaluate its effectiveness to determine whether it supports or harms the overall health of the United States.

Despite the policy's important goals, there have been a lot of disagreements over the actual effectiveness of the policy to achieve these goals. By comparing data before and after the policy, the study looks at how the policy has changed important indicators of child health and food security. We want to find out whether the HHFKA was effective in improving child health and food security and to determine where the policy might need to be refined.

⁸ https://obamawhitehouse.archives.gov/sites/default/files/Child_Nutrition_Fact_Sheet_12_10_10.pdf (Access Date: March 2025)

⁹ <https://www.fns.usda.gov/pd/child-nutrition-tables> (Access Date: March 2025)

The contribution of the paper can be identified as follows. The paper analyzes an extensive number of key variables that are essential indicators of the success of the policy through t-test and regression analysis. The variables, deviating from much of the previous research on this topic, are not only indicators of the effectiveness of the policy's implementation, like the quality of the food served, but also indicate the effectiveness of the policy's effects on the outcome of children's health and food security.

This paper proceeds as follows. Section 2 provides background information on the existing literature. Section 3 highlights the empirical findings, and Section 4 concludes the paper.

Literature Review

The Healthy, Hungry Free Kids Act (HHFKA) of 2010 was an important step in improving health outcomes and nutrition standards of school-aged children in the United States. The literature review summarizes findings from existing studies investigating the implementation and outcomes of the HHFKA using diverse metrics.

Several studies demonstrated significant improvements in the average nutritional quality of school lunches after the implementation of the HHFKA. Johnson et al. (2016) reported an increase in the Monthly Mean Adequacy Ratio (MAR) from a score of 58.72 to 75.6, and a decrease in Mean Energy Density (ED) from 1.65 to 1.44, highlighting improvements made in the nutritional quality of school lunches served. Additionally, Mozer et al. (2019) reported a similar improvement in MAR and energy per serving for all entrées combined, based on data from a Washington State school district from 2011 to 2014. Similarly, Gearan and Fox (2020) observed an increase in the Total Healthy Eating Index 2010 (HEI-2010) scores from 58% of the maximum score in the 2009-2010 school year to 82% of the maximum score in the 2014-2015 school year. In addition, Vaudrin et al. (2014) also found a significant improvement in the selected and consumed key nutrients, including reductions in sodium and the percentage of calories from saturated fat, and a significant increase in fiber. Furthermore, Kinderknecht et al. (2020) and Berger et al. (2020) corroborated through two different methods that there was an improvement in dietary quality for participants in the HHFKA, regardless of their family's income level.

The impact of the HHFKA on school lunch participation rates had been negligible. For example, Johnson et al. (2016) noted negligible changes in school meal participation rates, from 46% to 47% of the mean number of daily meals served per student enrolled, while Hecht et al. (2017) found that the HHFKA did not hurt school meal participation over time, causing only a minor increase from 70% to 72% of participation in the National School Lunch Program.

However, Asada et al. (2020) revealed gaps in awareness and implementation of USDA wellness policies among school superintendents, emphasizing the need for improved communication and support.

The HHFKA has also been associated with mixed effects on Body Mass Index (BMI) and obesity. For example, Kenny et al. (2020) found there was no significant association between the policy and childhood obesity trends overall but the risk of obesity for children in poverty

declined substantially each year after the act's implementation. Chandran et al. (2023) observed a significant decrease in BMI among school-aged youths in the US. Richardson et al. (2022) found that high BMI progression among low-income children was mitigated post-HHFKA.

Despite the positive outcome of the HHFKA, several studies identified challenges in implementation. Rida et al. (2019) and Weir and Sharma (2016) reported time constraints, limited administrative support, and clarity issues in the policy guidelines. Evans et al. (2021) found that faculty in middle schools felt schools needed more flexibility and decision-making power in nutrition reform for the specific needs of their community for more success.

Studies on food selection and waste reveal mixed outcomes. Cohen et al. (2014) observed improvements in meal selection and consumption despite the continued high levels of food waste. Cullen et al. (2015) found a similar increase in the selection and consumption rates for fruit, whole grains, and most vegetables after the policy.

The Healthy Hungry Free Kids Act represents a significant step in improving child nutrition and health, especially emphasized in low-income students. However, issues like administrative challenges and food waste remain. Providing more support to schools and making specific changes to the policy could increase the effectiveness and long-lasting impact of the policy.

Authors	Publication Year	Country /region Studied	Variable Used	Methods	Time Period of Data
Johnson DB et al.	2016	U.S.	Monthly mean adequacy ratio (MAR), mean energy density (ED), Monthly school meal participation	Descriptive and longitudinal Study	2011 – 2014
Kenney, Erica L., et al.	2020	U.S.	Child Obesity	Time series analysis	2003-2018
Kinderknecht et al.	2020	U.S.	Healthy Eating Index-2010	Cross-sectional study	2007-2010; 2013-2016
Gearan et al.	2020	U.S.	Healthy Eating Index 2010	Two cross-sectional analyses	2014-2015; 2009-2010
Richardson et al.	2022	U.S.	Body Mass Index	Cross-cohort comparison	1998-99; 2010-11
Chandran, et al.	2023	U.S.	Body Mass Index	Time-series analysis	2005-2020
Hecht, Amelie A., et al.	2017	U.S.	Annual average daily participation rates	Longitudinal analyses, linear models	2008–2009; 2014–2015
Vaudrin, et al.	2014	U.S.	Selected and consumed key nutrients	MANOVA	2012-2013
Mozer, et al.	2019	U.S.	Nutritional quality of entrée by mean adequacy ratio, energy density, and energy per serving.	Wilcoxon two-sample test	2011-2014
Au, Lauren E., et al.	2020	U.S.	Percent Adherence with federal nutrition standards	Cross-sectional analysis	2013-2015
Asada, Y., et al.	2020	U.S.	Superintendents' perspective themes	Qualitative analysis, Participant interviews,	2017
Rida, Z., et al.	2019	U.S.	Mixed-methods themes of school food-service professionals' perception	Mixed methods approach with qualitative and quantitative analysis	2012
Prokop, J. L., et al.	2011	U.S.	Obesity, Nurse support	Literature review	2011
Martinelli, S., et al.	2020	U.S.	Parental perception of school meals student participation	Multivariable logistic regression	2009–2010; 2016–2017
Evans, R. R., et al.	2021	U.S.	Middle School Stakeholders (faculty)	Qualitative analysis	2021
Weir, K., et al.	2016	U.S.	Child nutrition professionals' feedback	Qualitative analysis	2012
Berger, A. T., et al.	2020	U.S.	Proportion of energy intake from school foods, association between school food intake and dietary quality,	Survey regression	2007-2016
Terry-McElrath, Y. M., et al.	2015	U.S.	foods and beverages offered	Time series analysis	2011-2013
Cohen, J. F., et al.	2014	U.S.	Plate waste	Logistic regression and mixed model	2011-2012
Cullen, K. W., et al.	2015	U.S.	Elementary students' school lunches selected and consumed	Time series analysis	2011-2013

Empirical Findings

Data

Appendix 1 presents various variables sourced from the WDI: World Bank¹⁰, the Household Food Security in the United States in 2022, ERR-325 USDA¹¹, the Economic Research Service: U.S. DEPARTMENT OF AGRICULTURE¹², and the National Center for Health Statistics¹³.

The policy variable is the dummy variable, a binary variable taking a value of 0 before the implementation of the Healthy, Hunger-Free Kids Act of 2010, and a value of 1 after the implementation of the policy. Healthcare and food accessibility variables include the probability of dying among adolescents ages 10-19 years (POD), the probability of dying among adolescents ages 5-9 years (PD), households with food-insecure children (HFI), National School Lunch Program lunches served (NSLS), National School Lunch Program free or reduced-price lunches served in a fiscal year (FRLS), Obesity (OBS), and Overweight (OVW)

POD is the probability of dying between 10-19 years of age expressed per 1,000 adolescents before reaching age 20, if subject to age-specific mortality rates of the specified year. PD is the probability of dying between 5-9 years of age expressed per 1,000 children before reaching age 10, if subject to age-specific mortality rates of the specified year. HFI is the food security of households that children reside in expressed as a percentage of households. Different children may be affected differently by the food security of their households. NSLS provides the number of lunches expressed in millions served to children by the National School Lunch Program in nearly 100,000 public and private schools and residential childcare institutions. FRLS is the number of free or reduced-priced lunches expressed in millions served to children in households at or below 130 percent of the Federal poverty line or households between 130 and 185 percent of the Federal poverty line respectively. OBS is the percent of the population between ages 2-19 with a body mass index (BMI) at or above the 95th percentile from the sex-specific BMI-for-age 2000 CDC Growth Charts. OVW is the percent of the population between ages 2-19 with a body mass index (BMI) at or above the 85th percentile and below the 95th percentile from the sex-specific BMI-for-age 2000 CDC Growth Charts.

¹⁰ <https://data.worldbank.org/> (Access Date: March 2025)

¹¹ <https://www.ers.usda.gov/sites/default/files/laserfiche/publications/107703/ERR-325.pdf?v=43507> (Access Date: March 2025)

¹² <https://www.ers.usda.gov/> (Access Date: March 2025)

¹³ <https://www.cdc.gov/nchs/index.html> (Access Date: March 2025)

Descriptive Statistics

Table 1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
POD	33	4.124	0.882	3	5.7
PD	33	0.752	0.179	0.6	1.1
HFI	22	9.841	1.518	6.8	12.3
NSLS	26	4746.458	669.082	2154.8	5278.4
FRLS	26	67.588	11.002	56.8	98.9
OBS	10	16.74	1.553	13.9	19.3
OVW	10	15.34	0.902	14.2	16.6

Table 1 provides descriptive statistics for various key indicator variables. The probability of dying among adolescents ages 10-19 years (per 1,000) has 33 observations with a mean of 4.124 and a standard deviation of 0.882, ranging from a minimum of 3 to a maximum of 5.7. The probability of dying among children ages 5-9 years (per 1,000) has 33 observations with a mean of 0.752 and a standard deviation of 0.179, ranging from a minimum of 0.6 to a maximum of 1.1. Households with food-insecure children (in %) have 22 observations with a mean of 9.841 and a standard deviation of 1.518, ranging from a minimum of 6.8 to a maximum of 12.3. National School Lunch Program lunches served in a fiscal year (in Millions) has 26 observations with a mean of 4746.458 and a standard deviation of 669.082, ranging from a minimum of 2154.8 to a maximum of 5278.4. National School Lunch Program free or reduced-price lunches served in a fiscal year (in Millions) has 26 observations with a mean of 67.588 and a standard deviation of 11.002, ranging from a minimum of 56.8 to a maximum of 98.9. Obesity in the population between ages 2-19 (in %) has 10 observations with a mean of 16.74 and a standard deviation of 1.553, ranging from a minimum of 13.9 to a maximum of 19.3. Overweight in the population between ages 2-19 (in %) has 10 observations with a mean of 15.34 and a standard deviation of 0.902, ranging from a minimum of 14.2 to a maximum of 16.6.

Findings

Table 2. Two-sample t-test with equal variances for the Probability of dying among adolescents ages 10-19 years (per 1,000)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 2 Mean	dif	St Err	p-value
POD	20	13	4.640	3.331	1.309	.216	0.000

Table 2 presents the results of a two-sample t-test assuming equal variance, comparing the probability of dying between two groups of adolescents ages 10-19 years. Group 1 represents adolescents ages 10-19 before the implementation of the policy and Group 2 represents adolescents ages 10-19 after the implementation of the policy. Group 1 has a higher mean

probability of dying (4.640) compared to Group 2 (3.331). The difference in mean is 1.309 with a standard error of 0.216. The p-value is 0.000, indicating that the difference in probabilities of dying between the pre and post-policy groups for adolescents ages 10-19 years is statistically significant.

Figure 1. Line representation of Probability of dying among adolescents ages 10-19 years (per 1,000)

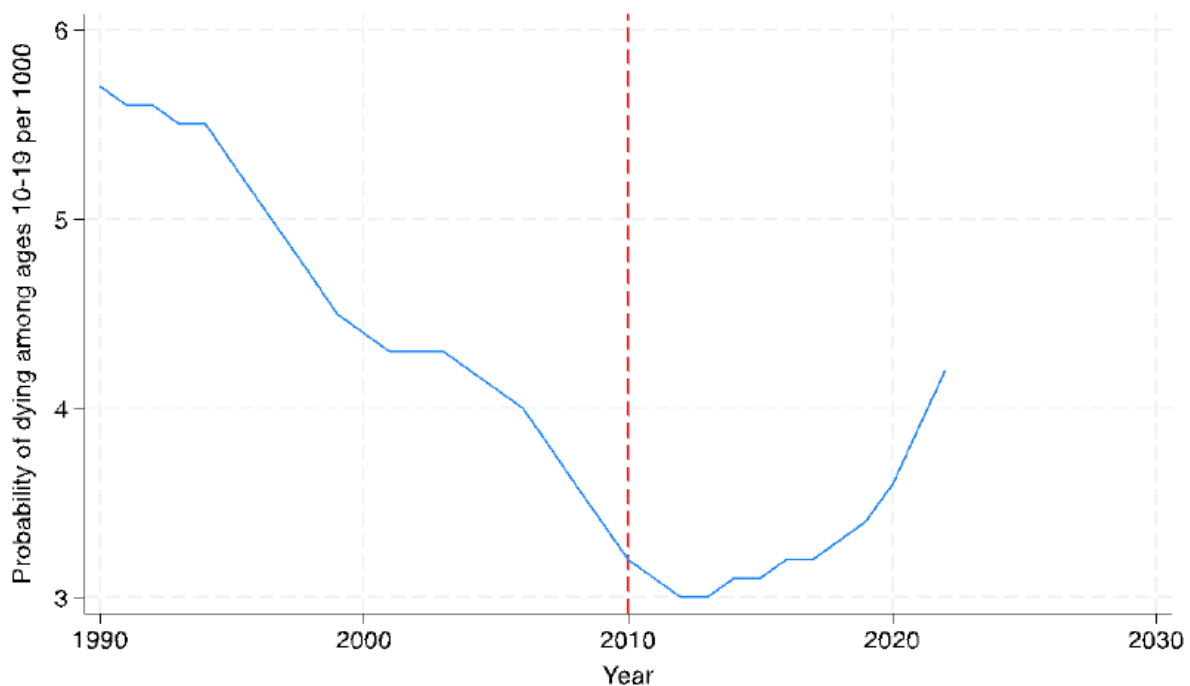


Figure 1 highlights that there was a significant increase in adolescent death rates after the policy was introduced. This suggests that the policy was not successful in reducing the probability of dying among adolescents ages 10-19 years. In other words, based on the line representation of the data, the policy had no additional effect on improving adolescent survival rates.

Table 3. Two-sample t-test with equal variances for Probability of dying among children ages 5-9 years (per 1,000)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 2 Mean	dif	St Err	p-value
PD	20	13	.85	.6	.25	.0465	0.000

Table 3 presents the results of a two-sample t-test assuming equal variance, comparing the probability of dying between two groups of children ages 5-9 years. Group 1 represents children ages 5-9 before the implementation of the policy and Group 2 represents children ages 5-9 after the implementation of the policy. Group 1 has a higher mean probability of dying (0.85) compared to Group 2 (0.60). The difference in mean is 0.25 with a standard error of 0.0465. The p-value is 0.000, indicating that the difference in probabilities of dying between the pre and

post-policy groups of children ages 5-9 years is statistically significant.

Figure 2. Line representation of Probability of dying among children ages 5-9 years (per 1,000)

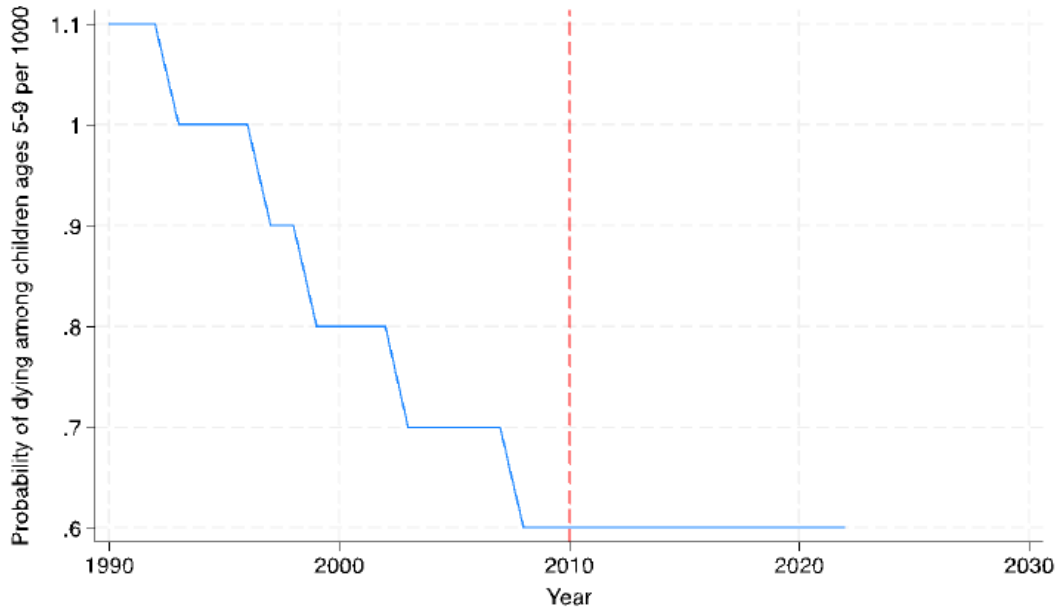


Figure 2 highlights that there was a significant difference in child death rates before and after the policy was introduced. This suggests that the policy was successful in reducing the probability of dying among children ages 5-9 years. In other words, the policy had a positive effect on improving child survival rates for children aged 5-9 years old.

Table 4. Two-sample t-test with equal variances for Households with food-insecure children (in %)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 1 Mean	dif	St Err	p-value
HFI	9	13	10.311	9.515	.795	.65	.236

Table 4 presents the results of a two-sample t-test assuming equal variance, comparing the percent of food-insecure children between two groups of households. Group 1 represents households before the implementation of the policy and Group 2 represents households after the implementation of the policy. Group 1 has a higher mean percentage (10.311) compared to Group 2 (9.515). The difference in mean is 0.795 with a standard error of 0.650. The p-value is 0.236, indicating that the difference in the percent of food-insecure children between the pre and post-policy groups of households is not statistically significant. Thus, the policy implementation has no significant effect on the food security of children.

Figure 3. Line representation of Households with food-insecure children (in %)

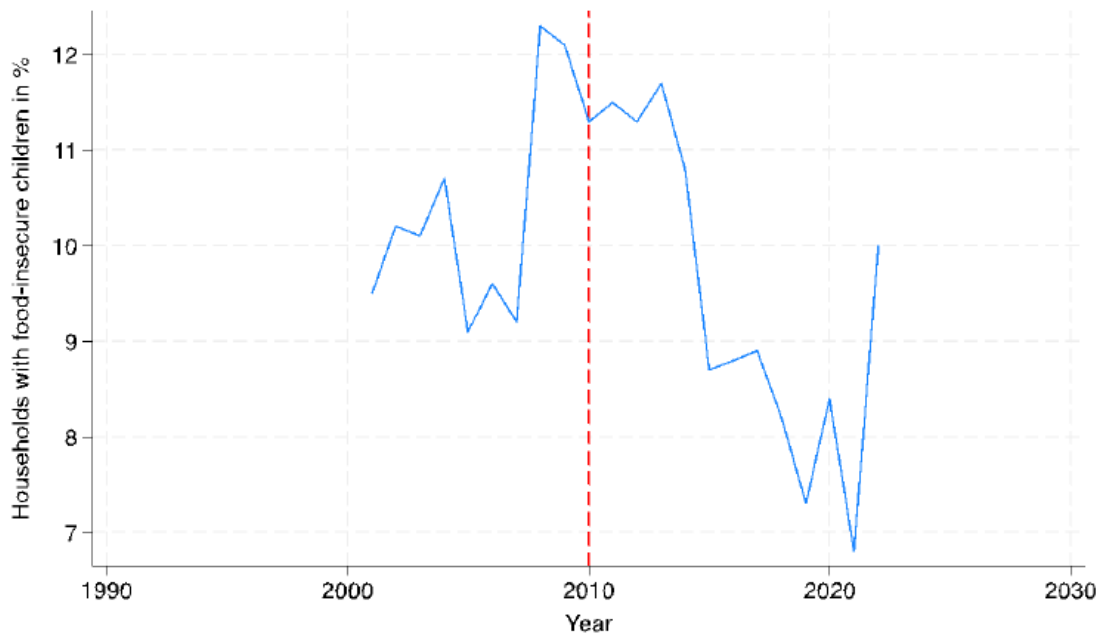


Figure 3 highlights that there was a difference in trend in households with food-insecure children before and after the policy was introduced. This suggests that the policy was successful in reducing the percentage of households with food-insecure children, however only by disregarding the year 2022. In other words, the policy had a positive effect on improving child food security. However, by accounting the year 2022 effects, there is not a significant difference between the percent households with food-insecure children before and after the policy. Since the t-test considers all years, including 2022, it describes no significant difference between the percentage of households with food-insecure children before and after the policy.

Table 5. Two-sample t-test with equal variances for National School Lunch Program lunches served in a fiscal year (in Millions)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 2 Mean	dif	St Err	p-value
NSLS	12	14	4824.234	4679.793	144.441	267.021	.594

Table 5 presents the results of a two-sample t-test assuming equal variance, comparing the number of lunches served in millions by the National School Lunch Program each fiscal year between two groups. Group 1 represents the number of lunches served before the implementation of the policy and Group 2 represents the number of lunches served after the implementation of the policy. Group 1 has a higher mean (4824.234) compared to Group 2 (4679.793). The difference in mean is 144.441 with a standard error of 267.021. The p-value is 0.594, indicating that the difference in the number of lunches served by the National School Lunch Program each

fiscal year between the pre and post-policy groups is not statistically significant. It is clear that policy implementation has no significant effect on the number of school lunches the NSLP served.

Figure 4. Line representation of National School Lunch Program lunches served in a fiscal year (in Millions)

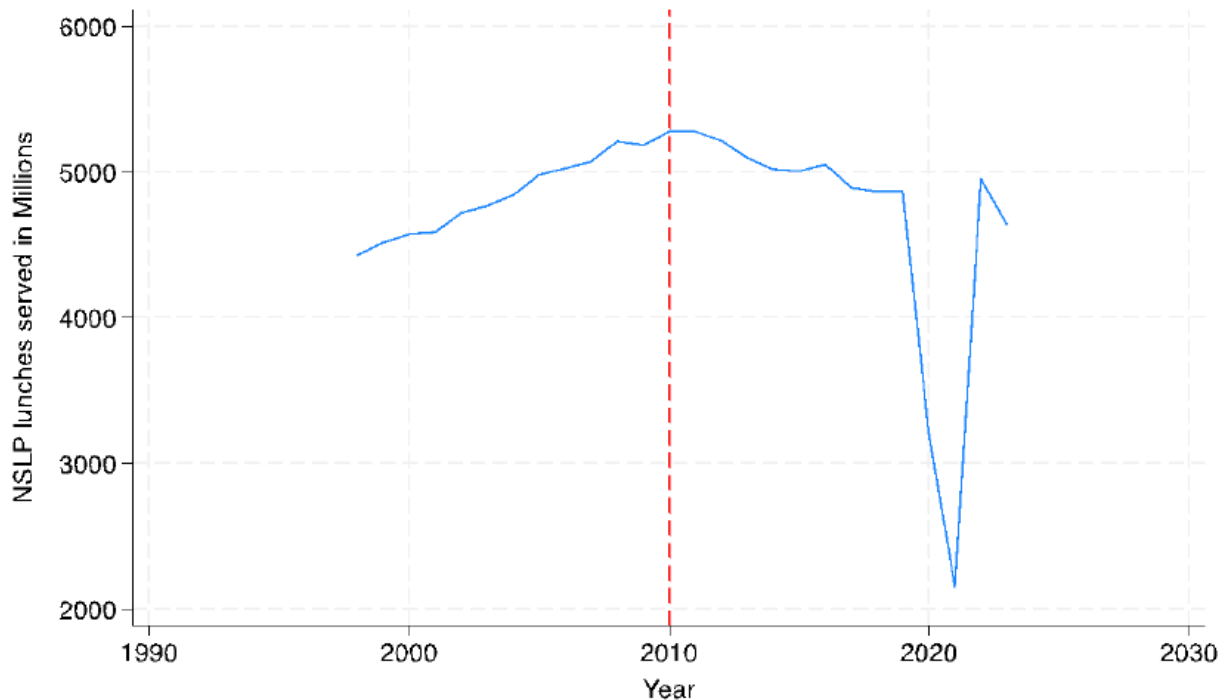


Figure 4 highlights that there was no difference in the number of lunches served by the National School Lunch Program before and after the policy was introduced. This suggests that the policy was not effective in increasing the number of lunches served, however only after disregarding the years 2020 and 2021. In other words, the policy did not affect improving the quantity of lunches served.

Table 6. Two-sample t-test with equal variances for National School Lunch Program free or reduced-price lunches served in a fiscal year (in Millions)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 2 Mean	dif	St Err	p-value
FRLS	12	14	58.767	75.15	-16.384	2.886	0.000

Table 6 presents the results of a two-sample t-test assuming equal variance, comparing the number of free and reduced-priced lunches in millions served by the National School Lunch Program each fiscal year between two groups. Group 1 represents the number of free and reduced-price lunches served before the implementation of the policy and Group 2 represents the number of free and reduced-price lunches served after the implementation of the policy. Group 1 has a higher mean (58.767) compared to Group 2 (75.15). The difference in mean is -16.384 with

a standard error of 2.886. The p-value is 0.000, indicating that the difference in number of free or reduced-priced lunches served by the National School Lunch Program each fiscal year between the pre and post-policy groups is statistically significant. It is clear that policy implementation has had a significant effect on the number of free or reduced-price school lunches the NSLP served.

Figure 5. Line representation of National School Lunch Program free or reduced-price lunches served in a fiscal year (in Millions)

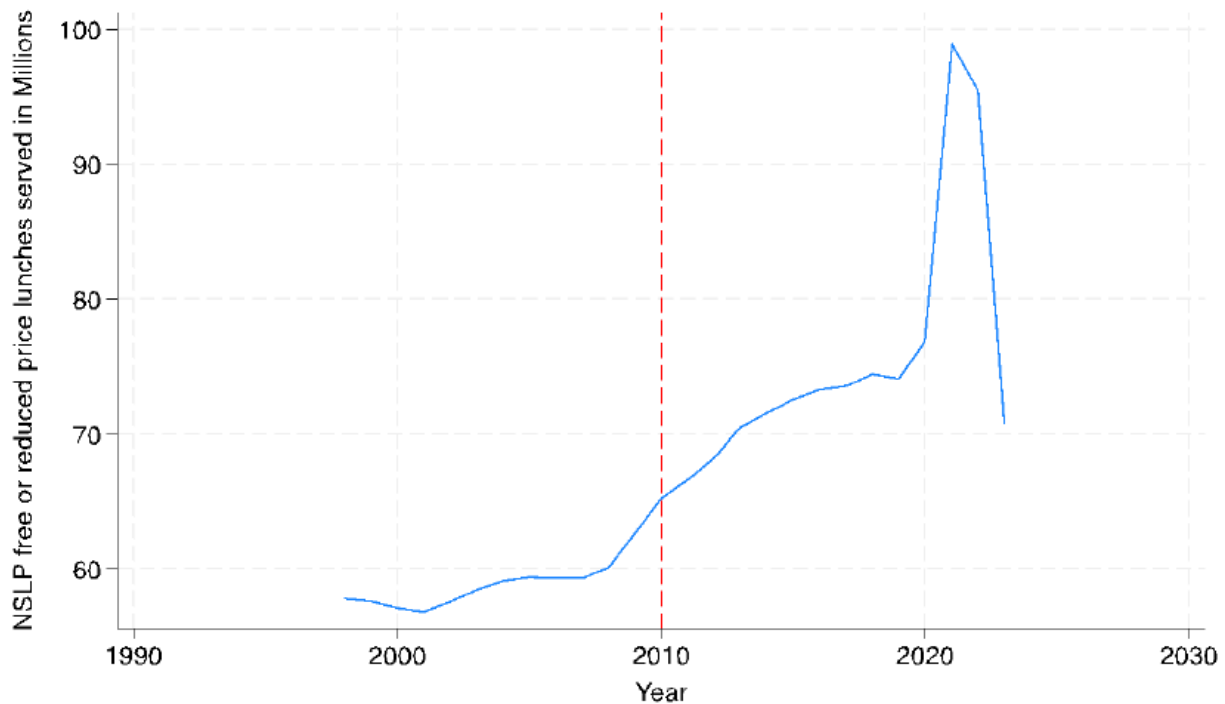


Figure 5 highlights that there was an increase in the number of free or reduced-priced lunches served by the National School Lunch Program before and after the policy was introduced. This suggests that the policy was successful in increasing the number of free or reduced-price lunches served. In other words, the policy had a positive effect on improving access to school lunches for students who need financial assistance.

Table 7. Two-sample t-test with equal variances for Obesity in population between ages 2-19 (in %)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 2 Mean	dif	St Err	p-value
OBS	6	4	15.917	17.975	-2.058	.775	.029

Table 7 presents the results of a two-sample t-test assuming equal variance, comparing the percent obese between two groups of people aged between 2 and 19. Group 1 represents the percent obese before the implementation of the policy and Group 2 represents the percent obese

after the implementation of the policy. Group 1 has a higher mean (15.917) compared to Group 2 (17.975). The difference in mean is -2.058 with a standard error of 0.775. The p-value is 0.029 indicating that the difference in percent obesity between the pre and post-policy groups of children aged 2-19 is statistically significant. It is clear that policy implementation has had a significant effect on the percentage obese of children aged 2-19.

Figure 6. Line representation of Obesity in the population between ages 2-19 (in %)

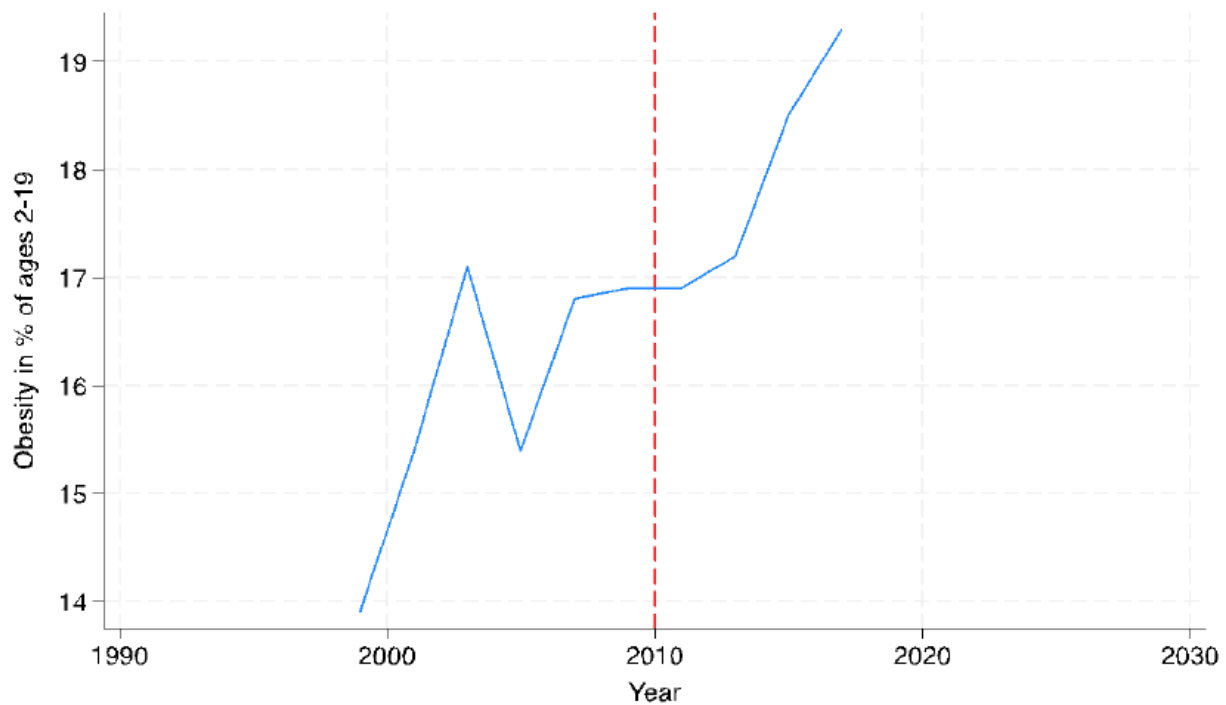


Figure 6 highlights that the percentage of obese children aged 2–19 increased from before to after the policy was introduced. This suggests that the policy was not successful in decreasing the percentage of obese children. In other words, the policy had a negative effect, instead of the intended positive one, of curbing obesity in children ages 2-19.

Table 8. Two-sample t-test with equal variances for Overweight in population between ages 2-19 (in %)

Variable Name	Grp 1	Grp 2	Grp 1 Mean	Grp 2 Mean	dif	St Err	p-value
OWV	6	4	14.934	15.95	-1.016	.503	.077

Table 8 presents the results of a two-sample t-test assuming equal variance, comparing the percent overweight between two groups of people aged between 2 and 19. Group 1 represents the percent overweight before the implementation of the policy and Group 2 represents the percent overweight after the implementation of the policy. Group 1 has a higher mean (14.934) compared to Group 2 (15.95). The difference in mean is -1.016 with a standard error of 0.503. The p-value is 0.077 indicating that the difference in percent overweight between the pre and

post-policy groups of children aged 2-19 is not statistically significant. Policy implementation has had a negligible effect on the percentage overweight of children aged 2-19.

Figure 7. Line representation of Overweight in population between ages 2-19 (in %)

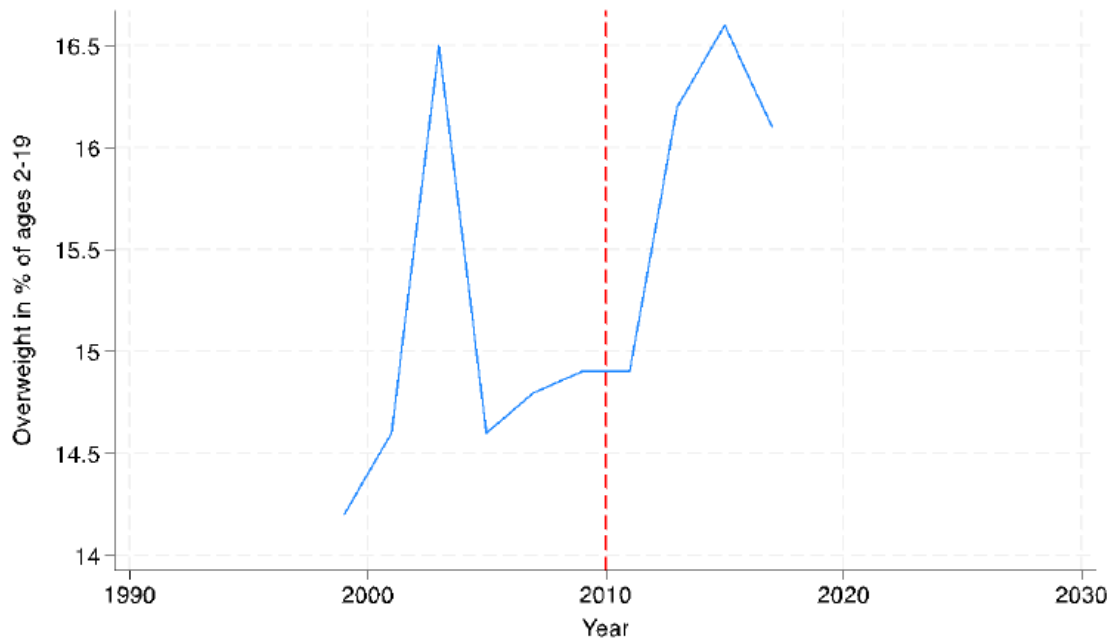


Figure 7 highlights that the percentage of overweight children aged 2–19 increased from before to after the policy was introduced. It rose to a number that had previously been reached in the mid-2000s. This suggests that the policy was not successful in decreasing the percentage of overweight children. In other words, the policy had a negligible or negative effect on curbing the percentage of overweight children ages 2-19.

Table 9. Regression Analysis

VARIABLES	(1) POD	(2) PD	(3) HFI	(4) NSLS	(5) FRLS	(6) OBV	(7) OVW
dummy	-1.3092*** (0.1900)	-0.2500*** (0.0376)	-0.7957 (0.6102)	-144.4405 (250.4016)	16.3833*** (2.6823)	2.0583** (0.7516)	1.0167* (0.4879)
Constant	4.6400*** (0.1616)	0.8500*** (0.0376)	10.3111*** (0.3905)	4,824.2333*** (77.3375)	58.7667*** (0.4584)	15.9167*** (0.5187)	14.9333*** (0.3351)
Observations	33	33	22	26	26	10	10
R-squared	0.5424	0.4816	0.0696	0.0120	0.5732	0.4685	0.3387

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 9 presents the result of 7 regression models, where the dependent variables are POD, PD, HFI, NSLS, FRLS, OBS, and OVW, respectively. The independent variable is the dummy variable, which is assigned a value of 1 when indicating the policy is in effect. This allows us to assess the impact of the policy on key indicators by comparing the outcomes before and after the implementation. In Model 1 and Model 2, the effect of the policy is negatively associated with POD and PD, with a coefficient of -1.31 and -0.250 respectively. The results are statistically significant at the 1% level, suggesting that the policy was successful in curbing the probability of dying among children in both ages 5-9 and ages 10-19. In models 5, 6, and 7, the effect of the policy is positively associated with FRLS, OBV, and OVW, with a coefficient of 16.38, 2.05, and 1.02, respectively. The results are statistically significant at the 1% level for FRLS, 5% level for OBV, and 10% level for OVW, suggesting that the policy was successful in providing free or reduced lunches to students, while unsuccessful in curbing obesity and overweight in children. We cannot capture a statistically significant effect of the policy on HFI and NSLS, indicating that the policy did not have a significant effect on the food security of children or the number of lunches served per fiscal year.

Conclusion

This paper aimed to evaluate the effectiveness of the Hungry Healthy Free Kids Act of 2010 in achieving its aims of improving children's health and food security. The paper analyzed seven key variables relevant to evaluating the effectiveness of the Act. Regression analysis, t-test analysis, and analysis of line representations of the data helped yield insightful results on these variables.

This paper's findings indicate mixed results in assessing the impact of the Hungry Healthy Free Kids Act of 2010. The results indicated a post-policy decrease in the probability of dying among the youth, indicating a possible positive effect on the Act towards child mortality. The results also indicated that the policy was successful in increasing meal access for low-income students as an increase in free or reduced-price lunches served was demonstrated. Furthermore, the results demonstrated that the more nutritious meals served after the implementation of the policy did not detract from the participation in the school lunch program. Despite this, the policy did not lead to significant improvement in reducing households with food-insecure children. Moreover, the policy didn't achieve its goal of reducing childhood obesity and overweight rates, highlighting the policy's ineffectiveness in curbing the obesity epidemic among the youth population in the United States. These results suggest that further adjustments in the policy are needed to improve the health of children in the long term and to ensure food security among the youth.

Works Cited

- Asada, Y., Harris, J. L., Mancini, S., Schwartz, M. B., & Chriqui, J. F. (2020). Food and beverage marketing in schools: school superintendents' perspectives and practices after the Healthy, Hunger-Free Kids Act. *Public Health Nutrition*, 23(11), 2024-2031.
- Au, L. E., Ritchie, L. D., Gurzo, K., Nhan, L. A., Woodward-Lopez, G., Kao, J., ... & Gosliner, W. (2020). Post-healthy, hunger-free kids act adherence to select school nutrition standards by region and poverty level: the healthy communities study. *Journal of nutrition education and behavior*, 52(3), 249-258.
- Berger, A. T., Widome, R., Erickson, D. J., Laska, M. N., & Harnack, L. J. (2020). Changes in association between school foods and child and adolescent dietary quality during implementation of the Healthy, Hunger-Free Kids Act of 2010. *Annals of epidemiology*, 47, 30-36.
- Bergman, E. A., Englund, T., Taylor, K. W., Watkins, T., Schepman, S., & Rushing, K. (2014). School lunch before and after implementation of the Healthy Hunger-Free Kids Act. *The Journal of Child Nutrition & Management*, 38(2), 1-12.
- Chandran, A., Burjak, M., Petimar, J., Hamra, G., Melough, M. M., Dunlop, A. L., ... & Knapp, E. (2023). Changes in body mass index among school-aged youths following implementation of the Healthy, Hunger-Free Kids Act of 2010. *JAMA pediatrics*, 177(4), 401-409.
- Child Nutrition Programs - National School Lunch Program | Economic Research Service.* (n.d.).
<https://www.ers.usda.gov/topics/food-nutrition-assistance/child-nutrition-programs/national-school-lunch-program>
- Cohen, J. F., Richardson, S., Parker, E., Catalano, P. J., & Rimm, E. B. (2014). Impact of the new US Department of Agriculture school meal standards on food selection, consumption, and waste. *American journal of preventive medicine*, 46(4), 388-394.
- Cullen, K. W., Chen, T. A., & Dave, J. M. (2015). Changes in foods selected and consumed after implementation of the new National School Lunch Program meal patterns in southeast Texas. *Preventive medicine reports*, 2, 440-443.
- Evans, R. R., Orihuela, C., & Mrug, S. (2021). Middle school stakeholder perceptions of school nutrition reform since the healthy, Hunger-Free Kids Act of 2010. *American Journal of Health Education*, 52(5), 276-287.
- Gearan, E. C., & Fox, M. K. (2020). Updated nutrition standards have significantly improved the nutritional quality of school lunches and breakfasts. *Journal of the Academy of Nutrition and Dietetics*, 120(3), 363-370.
- Johnson, D. B., Podrabsky, M., Rocha, A., & Otten, J. J. (2016). Effect of the Healthy Hunger-Free Kids Act on the nutritional quality of meals selected by students and school lunch participation rates. *JAMA pediatrics*, 170(1), e153918-e153918.
- Kenney, E. L., Barrett, J. L., Bleich, S. N., Ward, Z. J., Cradock, A. L., & Gortmaker, S. L. (2020). Impact Of The Healthy, Hunger-Free Kids Act On Obesity Trends: Study examines impact of the Healthy, Hunger-Free Kids Act of 2010 on childhood obesity trends. *Health Affairs*, 39(7), 1122-1129.

- Kinderknecht, K., Harris, C., & Jones-Smith, J. (2020). Association of the healthy, hunger-free kids act with dietary quality among children in the US national school lunch program. *Jama*, 324(4), 359-368.
- Martinelli, S., Acciai, F., Au, L. E., Yedidia, M. J., & Ohri-Vachaspati, P. (2020). Parental perceptions of the nutritional quality of school meals and student meal participation: before and after the healthy hunger-free kids act. *Journal of nutrition education and behavior*, 52(11), 1018-1025.
- Mozer, L., Johnson, D. B., Podrabsky, M., & Rocha, A. (2019). School lunch entrées before and after implementation of the Healthy, Hunger-Free Kids Act of 2010. *Journal of the Academy of Nutrition and Dietetics*, 119(3), 490-499.
- Products - Health e Stats - Prevalence of overweight, obesity, and severe obesity among children and adolescents aged 2–19 years: United States, 1963–1965 through 2017–2018.* (n.d.). <https://www.cdc.gov/nchs/data/hestat/obesity-child-17-18/obesity-child.htm>
- Prokop, J. L., & Galon, P. (2011). Healthy, Hunger-Free Kids Act of 2010: An opportunity for school nurses to make a difference. *The Journal of School Nursing*, 27(6), 401-403.
- Richardson, A. S., Weden, M. M., Cabrerros, I., & Datar, A. (2022). Association of the Healthy, Hunger-Free Kids Act of 2010 with body mass trajectories of children in low-income families. *JAMA Network Open*, 5(5), e2210480-e2210480.
- Rabbitt, M. P., Hales, L. J., Burke, M. P., Coleman-Jensen, A., & Economic Research Service. (2023). *Household food security in the United States in 2022* (No. 325). Economic Research Service. <https://www.ers.usda.gov/webdocs/publications/107703/err-325.pdf?v=5582.4>
- Rida, Z., Hall, E., Hasnin, S., Coffey, J., & Dev, D. A. (2019). The perception of school food-service professionals on the implementation of the Healthy, Hunger-Free Kids Act of 2010: a mixed-methods study. *Public health nutrition*, 22(11), 1960-1970.
- Terry-McElrath, Y. M., O'Malley, P. M., & Johnston, L. D. (2015). Foods and beverages offered in US public secondary schools through the National School Lunch Program from 2011–2013: early evidence of improved nutrition and reduced disparities. *Preventive medicine*, 78, 52-58.
- Vaudrin, N., Lloyd, K., Yedidia, M. J., Todd, M., & Ohri-Vachaspati, P. (2018). Impact of the 2010 US Healthy, Hunger-Free Kids Act on school breakfast and lunch participation rates between 2008 and 2015. *American journal of public health*, 108(1), 84-86.
- Weir, K., & Sharma, A. (2016). Child nutrition professionals' feedback on implementing the Healthy, Hunger-Free Kids Act of 2010. *Journal of Foodservice Business Research*, 19(2), 213-221.
- White House. (2010). Child nutrition reauthorization Healthy. *Hunger-Free Kids Act of*. *World Bank Open Data*. (n.d.). World Bank Open Data. <https://data.worldbank.org/>

Ketamine's Rapid-Acting Antidepressant Effects are Mediated in Part Through Anti-Inflammatory Mechanisms By Roger Li

ABSTRACT

Currently, available therapies for depression are still unable to treat a large portion of patients, while the well-known dissociative anesthetic ketamine has shown potential as a faster and more effective treatment. While ketamine's antidepressant mechanisms are still unknown, recent work has indicated that ketamine's anti-inflammatory properties play a critical role in its antidepressant effects. Ketamine modulates the inflammatory response through several distinct mechanisms, including metabolic, immune, and other pathways. Metabolically, ketamine modulates the kynurenine pathway, which is necessary for sustaining the inflammatory response. Ketamine has been shown to alter the microbiome, which dramatically affects the microbiota-gut-immune axis, and it also directly modulates immune pathways, including the autophagy-NLRP3 pathway, apoptosis, and pyroptosis. These many anti-inflammatory effects of ketamine support the long-hypothesized relationship between inflammation and depression, and provide insight into leveraging the inflammatory pathway in treatments for depression. A stronger understanding of ketamine's effects is vital to improve ketamine's antidepressant effects, reduce side effects, and inform future therapeutic treatments for depression.

INTRODUCTION

Depression as a disease

Depression, also known as major depressive disorder (MDD), affects 322 million people worldwide (Johnston et al., 2023). Approximately 40% of depression is heritable, which leaves room for many influencing factors, such as sex, environment, age, drug use, and socioeconomic status. Though the cause of depression is not yet clear, inflammation has long been implicated, as it has been correlated with many symptoms of depression. For example, depression is two to three more likely to occur in women, who also suffer from auto-immune and inflammatory disease at higher rates than men (Johnston et al., 2023). Furthermore, some symptoms of depression are similar to those caused by immune system activation, such as loss of energy, loss of appetite, and fatigue.

Depression is a complex and heterogeneous disease, meaning that it appears differently in each person. However, the two standards for diagnosing depression share nine symptoms, of which five must be met for a diagnosis: depressed mood; markedly diminished interest or pleasure in activities; reduced ability to think or concentrate, or indecisiveness; feelings of worthlessness, or excessive or inappropriate guilt; recurrent thoughts of death, or suicidal ideation, or suicide attempts or plans; insomnia or hypersomnia; significant change in appetite or weight; psychomotor agitation or retardation; and fatigue or loss of energy (Maj et al., 2020). The three main methods of treating depression are 1) antidepressants and other medications that cause antidepressant effects, 2) psychotherapy, and 3) somatic non-pharmacological treatments,

which are physical interventions that affect the brain by non-pharmaceutical means (Beurel et al., 2020). Despite the variety of treatment options, many people are resistant to treatment. In one study, treatment with either an antidepressant or psychotherapy resulted in remission in only 50% of patients (Dunlop et al., 2017). In the USA, it is accepted that approximately one-third of patients are non-respondent to treatment with traditional antidepressants, a condition known as treatment-resistant-depression (TRD) (Johnston et al., 2023). Thus, research into depression is still ongoing and necessary for the development of future therapeutic interventions.

Overview of the inflammatory response

Inflammation refers to the body's immune response, which is generally activated by irritation, injury, or infection. Common symptoms of inflammation include redness, swelling, and pain. The inflammatory response is mediated by a variety of cell types. The immune system is composed of the innate and adaptive immune systems. The innate immune system includes macrophages and dendritic cells, which, when activated, produce cytokines that recruit immune cells to the infected location. Cytokines are proteins that can enhance or reduce the inflammatory response. These include pro-inflammatory cytokines such as tumor necrosis factor (TNF- α), some interleukin forms (e.g. IL-1 β , IL-6), prostaglandins, or interferons. These cytokines and inflammatory factors not only amplify the immune response, but they also indicate the presence of inflammation. To transition to the adaptive immune response, dendritic cells present antigens—foreign molecules such as viruses—to activate lymphocytes. Lymphocytes are white blood cells that are a key component of the adaptive immune system and include B and T cells. Activated B cells become plasma cells, which produce antibodies that can neutralize antigens. Activated T cells become 1) cytotoxic T cells, which induce apoptosis 2) T helper cells, which influence activity of other immune cells 3) regulatory T cells, which regulate immune response to prevent autoimmunity. Some B and T cells become memory cells, which can facilitate a quicker response to the same pathogen in future. There also exist immune cells that are specific to the immune response of the brain, including microglia and astrocytes, which behave like innate immune cells. They can also be referred to collectively as glial cells, which are non-neuronal cells in the brain.

Correlations between inflammation and depression

Emerging research shows that inflammation may have a significant role in depression. A meta-analysis of individuals with depression found that 75% of participants had low-grade inflammation (CRP levels greater than 3 mg/L, the threshold for what is normal) (Johnston et al., 2023). Similarly, positron emission tomography (PET) scans of translocator protein 18 kDa (TSPO) – an inflammatory marker – have consistently shown greater TSPO binding in both the prefrontal cortex (PFC) and anterior cingulate cortex (ACC) areas of the brain in participants in the midst of a major depressive episode (Meyer et al., 2020). Both the elevation of proinflammatory factors and the decrease of anti-inflammatory factors are seen in both peripheral and cerebrospinal fluid and correlate with depressed patients' symptoms and disease duration (Guo et al., 2023). Additionally, the remission of depressive symptoms is accompanied by a decrease in inflammatory factors (Guo et al., 2023). Accordingly, a variety of antidepressants

result in the downregulation of a multitude of peripheral inflammatory markers and inhibit neuroinflammation (Guo et al., 2023). Overall, studies suggest that elevation of inflammation is correlated with more severe depression, higher recurrence of depressive symptoms, more adverse effects on brain function, such as impaired brain connectivity within motivation and reward circuits, and increased suicidality (Johnston et al., 2023). Though much research into the relationship between inflammation and depression is centered around inflammation within the brain, it is still viable to measure peripheral inflammation in this context since it correlates with central nervous system inflammation (Guo et al., 2023). Peripheral inflammatory markers can get into the brain by either crossing the blood-brain barrier through leakage zones of less strictness or by activating various cell types, such as cerebrovascular endothelial cells, which would be able to allow cytokines and other inflammatory factors into the brain (Beurel et al., 2020).

There are numerous studies supporting the correlation between depression and inflammation, while other studies exist that find no association between elevated levels of CRP and IL-6 (common inflammatory markers). Thus, the relationship between depression and inflammation is still unclear.

Role of Inflammation in Depression

Research suggests that inflammation may play a causal role in depression. One longitudinal study found a positive correlation between participants with elevated IL-6 and CRP levels at age nine and their likelihood of depression at age eighteen (Johnston et al., 2023). Additionally, a review found that cancer patients with hyperactivity of the HPA axis caused by cancer and anticancer treatments, which led to subsequent inflammation, exhibited increased depressive-like symptoms (Johnston et al., 2023).

One potential mediator between inflammation and depression is glutamate, a common excitatory neurotransmitter. Activated microglial cells may cause excessive release of glutamate, which leads to hyperactive signaling of glutamate receptors, which ultimately results in synaptic dysfunction and destruction. Additionally, glutamate diffusion away from the synapse contributes to loss of neurotransmission specificity, which can worsen depressive-like behaviors. Glutamine is also involved in the glutamate-glutamine cycle. Glutamine, the precursor to glutamate, may also play a role. One study found a correlation between depression severity, elevated levels of glutamine, and elevated glutamine-to-glutamate ratios (Johnston et al., 2023) in patients with depression. There have also been differences in cycle components that have been observed in postmortem brain tissue from healthy individuals and those with MDD (Johnston et al., 2023).

Not only are molecular biomarkers of depression affected by neuroinflammation, but depression symptoms are as well. Two common symptoms of depression are anhedonia and amotivation, and they have been consistently shown to be affected by inflammation levels. Inflammation likely disrupts dopamine activity in reward circuits, contributing to anhedonia (loss of pleasure) and reduced motivation (Johnston et al., 2023). Dopamine is a neurotransmitter associated with reward and motivation. Inflammation can dampen dopamine activity within reward circuits, resulting in disrupted frontostriatal functional connectivity (activity between

brain regions) and ultimately mediating motivational and behavioral responses (Johnston et al., 2023). In essence, inflammation can negatively influence the neural circuits underlying motivational symptoms closely related to anhedonia (Johnston et al., 2023). This is further supported by research showing a significant correlation between cytokine levels (central IL-6 soluble receptor (IL-6sr) and peripheral CRP levels) and reduced motivation and anhedonia (amongst other depressive symptoms) (Johnston et al., 2023).

Additionally, inflammation affects signaling in brain regions directly related to depression. One study of depressed patients specifically found a correlation between plasma CRP levels and decreased connectivity between the ventral striatum and ventromedial PFC (vmPFC) (Johnston et al., 2023). The change in connectivity between these areas has been shown to be correlated with the severity of anhedonia. A clinical study involving the administration of IFN- α , an interferon, for 4–6 weeks in fourteen individuals without depression both induced anhedonia and reduced bilateral activation, or activation on both hemispheres of the brain, of the ventral striatum in an fMRI reward task (Johnston et al., 2023). Lastly, reduced motivation has been correlated with levels of TNF- α within the central nervous system.

Depression can be treated with ketamine

Ketamine is an NMDA antagonist, meaning it blocks the NMDA receptor, which is involved in neuroplasticity. Originally licensed by the FDA for use as an anesthetic, ketamine exists in two forms: R-ketamine and S-ketamine. R-ketamine and S-ketamine are enantiomers—two different molecules that are mirror images of each other. S-ketamine, also known as esketamine, comprises the majority of ketamine research, especially in clinical trials (Johnston et al., 2023). It is FDA-approved for the treatment of depression as a nasal spray, though R-ketamine is not. However, R-ketamine has shown promise in early clinical trials since it may possess stronger antidepressant effects while also inducing fewer side effects (Johnston et al., 2023). Some ketamine metabolites (broken-down products of ketamine by metabolism) have also shown promise as antidepressants in preliminary research (Johnston et al., 2023).

Ketamine's response rate for patients with TRD is 25% to 85% at 24 hours after infusion (Johnston et al., 2023). This demonstrates two main benefits of ketamine over traditional antidepressants. Firstly, as previously mentioned, one-third of patients with depression do not respond to treatment with traditional antidepressants, so ketamine's ability to treat a large percentage of them is noteworthy. Secondly, traditional antidepressants often take weeks or months to take effect, whereas ketamine is generally able to show benefits within hours or days. Ketamine is also able to treat symptoms that are typically treatment-refractive under traditional antidepressants, such as anhedonia, amotivation, and suicidality (Johnston et al., 2023).

Ketamine's mechanisms of action for its antidepressant effects are still ambiguous and multi-faceted, but researchers primarily attribute ketamine's therapeutic effects to upregulated neuroplasticity induced via glutamatergic modulation (Johnston et al., 2023).

Neuroinflammation resulting from the Gut-Brain Axis

The gut-brain axis (GBA) refers to the bidirectional communication system between gut microbiota, or microbiome, and the brain. The gut refers to the gastrointestinal tract (GI), which includes organs such as the stomach and small and large intestines, and the gut microbiota refers to the trillions of bacteria that are present within the gut. This communication system uses three main pathways: neural (through neurons), inflammatory (through cytokines and immune cells), and humoral (through the HPA axis, consisting of the hypothalamus, pituitary glands, and adrenal glands). The gut-brain axis has been shown to be interrelated with a variety of mental disorders, including MDD, autism, and anxiety disorders (Wilkowska et al., 2021).

Relationship between Gut-Brain axis and inflammation

It has been suggested that the inflammation that is seen in depressed patients is a result of changes at the microbiota levels or at least correlated (Beurel et al., 2022). There is evidence of microbiome alterations in depressed patients, which suggests that they may contribute to dysregulated inflammatory responses (Rogers et al., 2016). Increased signaling in the NLRP3 inflammasome, a protein involved in the immune response, led to an increase in pro-inflammatory markers within gut-microbiota (Kouba et al., 2022). Another study showed that fecal microbiota transplantation in animals with depression induced by chronic unpredictable stress (CUS) was able to restore decreased levels of serotonin and suppress activity of microglia and astrocytes as well as reduce expression of NLRP3 and other pro-inflammatory cytokines within the hippocampus of the animals (Rao et al., 2021).

Apart from inflammation, the gut microbiome has been implicated in depression through its ability to increase BDNF expression in the brain. BDNF increases neurogenesis, which is the production of new neurons, and has been shown to be strongly decreased in depression (Rana 2020). In one study, the addition of two strains of bacteria into the gut microbiota was able to increase BDNF levels in patients with MDD and other psychiatric disorders (Dehghanai et al., 2022). Overall, the gut-brain axis holds potential as a pathway to mediating inflammation and depression.

METHODS

Research methods involving animals

Animals used included rats and mice. Many animal models are used to induce depressive-like symptoms such as anhedonia (inability to feel pleasure) and learned helplessness (lack of motivation to change unpleasant circumstances). Animals subjected to these models will retain depressive-like symptoms after the termination of the stressor, allowing for chronic antidepressant research. *In vitro* experiments with microglia from animals are also used (Lyu et al., 2022).

One animal model is the chronic restraint stress-induced depression-like model (Lyu et al., 2022). It involves restraining the rats into constrained cages in other devices. It is done for multiple hours a day and up to multiple weeks (Lyu et al., 2022). Another model is chronic unpredictable stress (CUS) or chronic mild stress (CMS). It involves continuous exposure to a

variety of mild stressors, such as periods of food and water deprivation, cage tilt, etc. (Dion-Albert et al., 2022). A very popular animal model for depression is the LPS inflammation-induced depression model. LPS refers to lipopolysaccharide, an antigen derived from bacteria that is able to activate the immune system. It is typically injected into mice intraperitoneally (through the abdomen). It can also be given by an infusion pump directly through a hole bored in the skull (Zhou et al., 2023). In 1996, it was also discovered to have use as a depression model because rats that were systemically injected with LPS displayed significantly decreased saccharin preference, indicating anhedonia (Yirmiya 1996).

There are several tests used that measure different depressive-like symptoms in animals. The sucrose preference test (SPT) involves depriving mice of water for 24 hours (Zhou et al., 2023). The animals are then given regular water and water with 2% sucrose. After being allowed to drink either beverage for a 2-hour (or other time) period, their sucrose preference is measured as a percentage of sucrose-water intake out of total water consumption (Zhou et al., 2023). This is used to measure anhedonia. The tail suspension test involves the suspension of animals by their tail. Their immobile time is measured. Higher immobile time correlates to greater levels of learned helplessness, another symptom of depression. The forced swim test (FST) involves placing mice in relatively deep water chambers (Zhou et al., 2023). They will naturally try to swim, and the time spent immobile is measured. Higher immobile time correlates to greater levels of learned helplessness. The open field test involves placing animals in a dimly lit environment and time spent exploring is measured (Lyu et al., 2022). Less time spent exploring correlates with higher depressive-like symptoms. Similarly, the elevated plus maze displays an extruded metal foundation. The outer sides of two opposite arms are closed, and the time that the mice spend on the open or closed arms are measured.

To measure the inflammatory response in these models of depression, protein expression of inflammation markers such as cytokines are measured in blood serum or cerebrospinal fluid (Lyu et al., 2022). Blood serum refers to blood plasma, blood without red or white blood cells, without clotting factors. Methods of measuring protein expression include ELISA kits, Western blot, immunohistochemistry, and immunofluorescence (Lyu et al., 2022).

Research methods involving humans

Most research surrounding ketamine's inflammatory effects on depression are done with animals. Humans are typically used when assessing the correlations between inflammation and depression.

Inflammatory markers are measured in blood plasma or blood serum using the same processes as with animals.

Neuroimaging can be used to measure inflammation or brain activity. To measure neuroinflammation, positron emission tomography (PET) scans are used to visualize inflammation distribution and intensity (Meyer et al., 2020). Functional connectivity, the activity between brain areas, is measured with fMRI (Costi et al., 2021).

To measure psychological outcomes, researchers can use a number of rating scales of depression including the Montgomery-Åsberg Depression Rating Scale (MADRS) (Chen et al., 2018).

RESULTS

Ketamine's effects on inflammation levels in animals

Ketamine can effectively decrease elevated levels of inflammatory markers in both animal models of depression and patients with depression. One in vitro study showed that, in glial cells with LPS-induced inflammatory responses, ketamine was able to reduce the inflammatory markers TNF- α and prostaglandin E(2) (Shibakawa et al., 2005). In another study involving LPS-activated macrophages, ketamine was able to decrease elevated levels of a number of proinflammatory markers, including TNF- α and IL-1 β (Tan et al., 2015).

In rodents, ketamine decreased cytokine elevation in blood plasma and various sections of the CNS, including the PFC, hippocampus, and spinal cord, while also decreasing microglial activation (Edem et al., 2023). Elevated levels of Th17 cells, a type of proinflammatory T helper cell, as well as an imbalance of the ratio between Th17 and T-regulatory cells, have both been associated with depression (Cui et al., 2021). Ketamine was able to inhibit differentiation and reactivation—the magnitude of cytokine production in already-developed immune cells—of Th17 cells (Lee, Jeong-Eun et al., 2017).

Ketamine's effects on inflammation levels in humans

In a study involving patients with TRD, a single intravenous dose of ketamine significantly reduced the pro-inflammatory cytokines IL-6 and IL- α four hours after injection (Kiraly et al., 2017). Another study also involving patients with TRD showed that ketamine induced rapid decreases of IL-6 and TNF- α elevations (Chen et al., 2018). It also found that greater decreases in TNF- α levels from before to after injection correlated with greater reduction in the MADRS score, reflecting an attenuation of depressive symptoms (Chen et al., 2018).

Surgery elicits a strong inflammatory response due to the tissue destruction involved. Ketamine, being used primarily as an anesthetic, has been studied in the surgical context for its anti-inflammatory properties. In patients undergoing cardiac surgery under extracorporeal circulation (ECC) (a situation that especially activates the immune system), an IV dose of ketamine as an anesthetic significantly reduced the expression of IL-6 (Roytblat et al., 1998). Notably, the reduction lasted for seven days after the operation, even though ketamine was out of the patient's system by that point. In liver transplant subjects, ketamine significantly reduced levels of TNF- α and IL-6 (Yang et al., 2006). A meta-analysis regarding the effects of ketamine on IL-6 expression in patients after surgery concluded that ketamine, given as an anesthetic before surgery, significantly offset the initial postoperative increase in IL-6 levels (Dale et al., 2012). This effect is comparatively as pronounced as the effects of methylprednisolone, one of the most potent anti-inflammatory medications, in cardiac surgery (Inaba et al., 1994).

Kynurenine pathway as a mediator of ketamine's antidepressant effects

The kynurenine pathway involves the metabolism of tryptophan into kynurenine. This pathway is controlled by indoleamine-2,3-dioxygenase (IDO), a rate-limiting enzyme. Upregulation of IDO, which can be caused by proinflammatory cytokines, increases the activity of the kynurenine pathway (Johnston et al., 2023). Tryptophan is competitively consumed as a precursor in the synthesis of serotonin, a neurotransmitter involved in mood regulation. Thus, proinflammatory cytokines indirectly reduce serotonin levels, which have been historically implicated in MDD. In support of the relevance of the pathway to MDD, there are higher ratios of kynurenine to tryptophan in patients with MDD, indicating kynurenine pathway activity since kynurenine is a product of tryptophan (Moaddel et al., 2018).

The kynurenine pathway has been implicated in neurogenesis. Hippocampal (HPC) neurogenesis, the generation of new neurons within a subsection of the hippocampus known as the dentate gyrus, mediates antidepressant effect (Santarelli et al., 2003). Cytokine-induced reductions in HPC neurogenesis are prevented by antidepressants (Borsini et al., 2017). This effect is mediated by decreased expression of kynurenine pathway components (Borsini et al., 2017).

Ketamine is able to prevent IDO expression and kynurenine production induced by pro-inflammatory cytokine IL-1 β . (Mandal et al., 2024). Additionally, the kynurenine pathway was shown to mediate the detrimental effects of IL-1 β on neurogenesis (Mandal et al., 2024). Thus, ketamine indirectly improved neurogenesis, a lack of which has been implicated in MDD.

Autophagy-NLRP3 pathway as a mediator of ketamine's antidepressant effects

Autophagy is the process of cellular recycling of unnecessary intracellular contents. It is implicated in the immune response through its connection to the NLRP3 inflammasome. Inflammasomes are protein complexes that are a part of the immune system; their function is to release cytokines. The NLRP3 inflammasome pathway starts by activation of the NLRP3 inflammasome, which leads to the activation of caspase-1, an enzyme. The enzyme cleaves immature pro-inflammatory cytokines, which mature into an active state. A lack of sufficient autophagy can induce the activation of inflammasomes, including the activation of the NLRP3 inflammasome in microglia (Kesharwani et al., 2019).

Ketamine activates autophagy, which in turn inhibits NLRP3 activity (Lyu et al., 2022). In a CRS model, ketamine was able to exert antidepressant effects (Lyu et al., 2022). This correlated with reductions in microglia activity and a reduction in inflammatory markers in the PFC and HPC, supporting the notion that inflammation is correlated with depression (Lyu et al., 2022). The decrease in inflammation was likely mediated via autophagy since autophagy increased after ketamine treatment (Lyu et al., 2022). Additionally, in the CRS model, when Baf-1, an autophagy inhibitor, was given alongside ketamine administration, ketamine's antidepressant effects were blocked.

In contrast, autophagy has also been cited as a process that ketamine inhibits for its antidepressant properties (Jiang et al., 2022). An LPS model, alongside inducing neuroinflammation and depressive-like symptoms, increased autophagy (Jiang et al., 2022). Ketamine was able to significantly reduce the adverse effects of LPS, including reducing cytokine levels, decreasing depression-like behavior, and decreasing autophagic markers (Jiang et al., 2022). This time, the use of autophagy inhibitor 3-MA did not alter ketamine's effects (Jiang et al., 2022). These disparate results have not been fully explained; however, it is possible that these results are not entirely contradictory. Rather, in low-grade inflammatory contexts, such as in chronic depression, autophagy activation may be necessary to clear inflammation-causing intracellular debris. However, when inflammation is very high in the brain (such as following LPS injection), ketamine's other anti-inflammatory effects may dampen inflammation enough to decrease the need for autophagy, thus resulting in lower levels of autophagic markers.

Apoptosis and pyroptosis pathways as mediators of ketamine's antidepressant effects

Apoptosis is controlled cell death; it is present in the immune system to get rid of infected cells. Ketamine has been suggested to both increase and decrease apoptosis. One *in vitro* study reported that ketamine reduced apoptosis. Additionally, ketamine decreases T cell apoptosis (Braun et al., 2010). In contrast, in an LPS model study, ketamine promoted autophagy, causing inhibition of the inflammatory response, thereby reducing apoptosis (Cao et al., 2022).

Pyroptosis is similar to apoptosis, but it actively promotes inflammation by releasing pro-inflammatory cytokines from ruptured cells. It is triggered by the activation of caspase-1, which is itself activated via the NLRP3 inflammasome. By decreasing the activation of NLRP3 (Lyu et al., 2022), ketamine would be expected to decrease pyroptosis. This was shown to be true in a chronic unpredictable mild stress model in which ketamine was able to decrease astrocyte pyroptosis in the HPC (Wen et al., 2024).

Gut-brain axis as a mediator of ketamine's antidepressant effects

The gut-brain axis and microbiome is linked to inflammation. Furthermore, increased NLRP3 signaling is commonly seen in patients with MDD, and this increases the amount of inflammatory markers within gut-microbiota (Kouba et al., 2022). Fecal microbiota transplantation suppressed the activity of microglia and astrocytes and reduced expression of NLRP3 and pro-inflammatory cytokine IL-1b in the HPC of chronic unpredictable stress (CUS)-exposed animals (Rao 2021), indicating that this axis plays a significant role in depression.

The antidepressant effects of ketamine may be mediated by its alteration of gut microbiota composition. Butyrate is a short-chain fatty acid—produced by bacterial fermentation in the gut—that can reduce inflammation. Ketamine administration increases butyrate producing microbiota including *Butyricimonas*, *Clostridiales*, *Turicibacter* and *Bacillus* (Getachew et al., 2018). It also reduces several kinds of bacteria that contribute to inflammation (Getachew et al., 2018; Liu et al., 2019).

Discussion

In this paper, we discussed multiple inflammatory pathways which ketamine may modulate to mediate its antidepressant effects. Given that each process has reasonable evidence supporting its relevance, it is unlikely that ketamine's antidepressant effects stem from a single anti-inflammatory process. Rather, the immune system as a whole has a significant role in MDD, and ketamine's disparate effects on it contribute to its antidepressant effects.. The inflammatory response and ketamine deserve further research because they may provide crucial answers in combating MDD.

Further questions about inflammation

More research should be done to improve specificity and clarify inconsistencies in results by controlling for more factors. Research should be done to clarify how the role of the immune system in depression varies by patients and if ketamine can have separate anti-inflammatory effects based on gender or other risk factors for depression.

Given the reliable evidence about inflammation's relationship to depression, an index or algorithm could be developed to predict the probability of developing MDD or responding to a certain drug. Additionally, questions remain about how non-pharmacological treatments influence inflammation. More research on inflammation is also warranted, including how different stressors increase the immune response and what factors determine its magnitude.

Further questions about ketamine

The majority of studies discussed here involved acute ketamine administration and its immediate effects. Future studies should also study the chronic administration and long-term effects in both preclinical and clinical settings. Additional research should be done to see how ketamine's anti-inflammatory and antidepressant effects can be prolonged.

More study is needed to learn the effects of R-ketamine, the lesser-studied enantiomer, and of ketamine metabolites. They each have differences in anti-inflammatory and antidepressant effects, which may offer insight into ketamine's mechanisms of action.

Ketamine has been shown to have side effects that should be considered before administration, including the potential for long-term brain cell damage, ketamine-related cystitis, and reduced cell viability (Chen et al., 2020; Yu et al., 2021; Yu et al., 2021). Further research into ketamine may be able to alleviate these effects.

While these and other questions remain about ketamine, its role in inflammation, and its mechanisms as an antidepressant, the drug has shown immense promise as a treatment for depression, particularly treatment-resistant depression. Further study into how ketamine modulates the immune system and ameliorates depressive symptoms will help develop ketamine into a safe and effective treatment for depression. Furthermore, this research will provide valuable insights into the mechanisms of inflammation and depression, paving the way for future targeted therapies for depression and related illnesses.

Works Cited

- Abelaira, Helena M et al. "The role of mTOR in depression and antidepressant responses." *Life sciences* vol. 101,1-2 (2014): 10-4. doi:10.1016/j.lfs.2014.02.014
- Amasi-Hartoonian, Nare et al. "Cause or consequence? Understanding the role of cortisol in the increased inflammation observed in depression." *Current opinion in endocrine and metabolic research* vol. 24 (2022): 100356. doi:10.1016/j.coemr.2022.100356
- Beurel, Eléonore et al. "The Bidirectional Relationship of Depression and Inflammation: Double Trouble." *Neuron* vol. 107,2 (2020): 234-256. doi:10.1016/j.neuron.2020.06.002
- Borsini, Alessandra et al. "Rescue of IL-1 β -induced reduction of human neurogenesis by omega-3 fatty acids and antidepressants." *Brain, behavior, and immunity* vol. 65 (2017): 230-238. doi:10.1016/j.bbi.2017.05.006
- Braun, S et al. "Ketamine induces apoptosis via the mitochondrial pathway in human lymphocytes and neuronal cells." *British journal of anaesthesia* vol. 105,3 (2010): 347-54. doi:10.1093/bja/aeq169
- Cao, Jianwei et al. "Ketamine Promotes LPS-Induced Pulmonary Autophagy and Reduces Apoptosis through the AMPK/mTOR Pathway." *Contrast media & molecular imaging* vol. 2022 8713701. 8 Jul. 2022, doi:10.1155/2022/8713701
- Chen, Mu-Hong et al. "Rapid inflammation modulation and antidepressant efficacy of a low-dose ketamine infusion in treatment-resistant depression: A randomized, double-blind control study." *Psychiatry research* vol. 269 (2018): 207-211. doi:10.1016/j.psychres.2018.08.078
- Chen, Yanni et al. "Yes-associated protein protects and rescues SH-SY5Y cells from ketamine-induced apoptosis." *Molecular medicine reports* vol. 22,3 (2020): 2342-2350. doi:10.3892/mmr.2020.11328
- Costi, Sara et al. "Peripheral immune cell reactivity and neural response to reward in patients with depression and anhedonia." *Translational psychiatry* vol. 11,1 565. 5 Nov. 2021, doi:10.1038/s41398-021-01668-1
- Cui, Meiying et al. "Th17 Cells in Depression: Are They Crucial for the Antidepressant Effect of Ketamine?." *Frontiers in pharmacology* vol. 12 649144. 15 Apr. 2021, doi:10.3389/fphar.2021.649144
- Dale, Ola et al. "Does intraoperative ketamine attenuate inflammatory reactivity following surgery? A systematic review and meta-analysis." *Anesthesia and analgesia* vol. 115,4 (2012): 934-43. doi:10.1213/ANE.0b013e3182662e30
- Dion-Albert, Laurence et al. "Sex differences in the blood-brain barrier: Implications for mental health." *Frontiers in neuroendocrinology* vol. 65 (2022): 100989. doi:10.1016/j.yfrne.2022.100989
- Dunlop, Boadie W et al. "Effects of Patient Preferences on Outcomes in the Predictors of Remission in Depression to Individual and Combined Treatments (PReDICT) Study." *The American journal of psychiatry* vol. 174,6 (2017): 546-556. doi:10.1176/appi.ajp.2016.16050517

- Dutton, Megan et al. "Stress, mental disorder and ketamine as a novel, rapid acting treatment." *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* vol. 65 (2022): 15-29. doi:10.1016/j.euroneuro.2022.09.006
- Edem, Edem Ekpenyong et al. "Ketamine abrogates sensorimotor deficits and cytokine dysregulation in a chronic unpredictable mild stress model of depression." *Psychopharmacology* vol. 239,1 (2022): 185-200. doi:10.1007/s00213-021-06021-4
- Getachew, Bruk et al. "Ketamine interactions with gut-microbiota in rats: relevance to its antidepressant and anti-inflammatory properties." *BMC microbiology* vol. 18,1 222. 22 Dec. 2018, doi:10.1186/s12866-018-1373-7
- Guo, Bingqi et al. "Neuroinflammation mechanisms of neuromodulation therapies for anxiety and depression." *Translational psychiatry* vol. 13,1 5. 9 Jan. 2023, doi:10.1038/s41398-022-02297-y
- Haroon, Ebrahim et al. "Antidepressant treatment resistance is associated with increased inflammatory markers in patients with major depressive disorder." *Psychoneuroendocrinology* vol. 95 (2018): 43-49. doi:10.1016/j.psyneuen.2018.05.026
- Haroon, Ebrahim et al. "Inflammation, Glutamate, and Glia: A Trio of Trouble in Mood Disorders." *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* vol. 42,1 (2017): 193-215. doi:10.1038/npp.2016.199
- Hua, Hao et al. "Depression and antidepressant effects of ketamine and its metabolites: The pivotal role of gut microbiota." *Neuropharmacology* vol. 220 (2022): 109272. doi:10.1016/j.neuropharm.2022.109272
- Inaba, H et al. "Suppression by methylprednisolone of augmented plasma endotoxin-like activity and interleukin-6 during cardiopulmonary bypass." *British journal of anaesthesia* vol. 72,3 (1994): 348-50. doi:10.1093/bja/72.3.348
- Jiang, Guanghao et al. "Autophagy: A New Mechanism for Esketamine as a Depression Therapeutic." *Neuroscience* vol. 498 (2022): 214-223. doi:10.1016/j.neuroscience.2022.05.014
- Johnston, Jenessa N et al. "Inflammation, stress and depression: An exploration of ketamine's therapeutic profile." *Drug discovery today* vol. 28,4 (2023): 103518. doi:10.1016/j.drudis.2023.103518
- Kesharwani, Radhika et al. "Interplay between Mitophagy and Inflammasomes in Neurological Disorders." *ACS chemical neuroscience* vol. 10,5 (2019): 2195-2208. doi:10.1021/acscchemneuro.9b00117
- Kiraly, D D et al. "Altered peripheral immune profiles in treatment-resistant depression: response to ketamine and prediction of treatment outcome." *Translational psychiatry* vol. 7,3 e1065. 21 Mar. 2017, doi:10.1038/tp.2017.31
- Kopra, Emma et al. "Ketamine's effect on inflammation and kynurenine pathway in depression: A systematic review." *Journal of psychopharmacology (Oxford, England)* vol. 35,8 (2021): 934-945. doi:10.1177/02698811211026426

- Lee, Jeong-Eun et al. "Inhibition of autoimmune Th17 cell responses by pain killer ketamine." *Oncotarget* vol. 8,52 89475-89485. 31 May. 2017, doi:10.18632/oncotarget.18324
- Li, Jia-Mei et al. "Ketamine may exert antidepressant effects via suppressing NLRP3 inflammasome to upregulate AMPA receptors." *Neuropharmacology* vol. 146 (2019): 149-153. doi:10.1016/j.neuropharm.2018.11.022
- Lyu, Dongbin et al. "Ketamine induces rapid antidepressant effects via the autophagy-NLRP3 inflammasome pathway." *Psychopharmacology* vol. 239,10 (2022): 3201-3212. doi:10.1007/s00213-022-06201-w
- Maj, Mario et al. "The clinical characterization of the adult patient with depression aimed at personalization of management." *World psychiatry : official journal of the World Psychiatric Association (WPA)* vol. 19,3 (2020): 269-293. doi:10.1002/wps.20771
- Mandal, Gargi et al. "Ketamine Prevents Inflammation-Induced Reduction of Human Hippocampal Neurogenesis via Inhibiting the Production of Neurotoxic Metabolites of the Kynurenine Pathway." *The international journal of neuropsychopharmacology* vol. 27,10 (2024): pyae041. doi:10.1093/ijnp/pyae041
- Meyer, Jeffrey H et al. "Neuroinflammation in psychiatric disorders: PET imaging and promising new targets." *The lancet. Psychiatry* vol. 7,12 (2020): 1064-1074. doi:10.1016/S2215-0366(20)30255-8
- Moaddel, Ruin et al. "Plasma metabolomic profiling of a ketamine and placebo crossover trial of major depressive disorder and healthy control subjects." *Psychopharmacology* vol. 235,10 (2018): 3017-3030. doi:10.1007/s00213-018-4992-7
- Rana, Tarapati et al. "Unfolding the Role of BDNF as a Biomarker for Treatment of Depression." *Journal of molecular neuroscience : MN* vol. 71,10 (2021): 2008-2021. doi:10.1007/s12031-020-01754-x
- Rao, Jingjing et al. "Fecal microbiota transplantation ameliorates stress-induced depression-like behaviors associated with the inhibition of glial and NLRP3 inflammasome in rat brain." *Journal of psychiatric research* vol. 137 (2021): 147-157. doi:10.1016/j.jpsychires.2021.02.057
- Roytblat, L et al. "Ketamine attenuates the interleukin-6 response after cardiopulmonary bypass." *Anesthesia and analgesia* vol. 87,2 (1998): 266-71. doi:10.1097/00000539-199808000-00006
- Shibakawa, Y Saito et al. "Effects of ketamine and propofol on inflammatory responses of primary glial cell cultures stimulated with lipopolysaccharide." *British journal of anaesthesia* vol. 95,6 (2005): 803-10. doi:10.1093/bja/aei256
- Tan, Yonghong et al. "Ketamine reduces LPS-induced HMGB1 via activation of the Nrf2/HO-1 pathway and NF- κ B suppression." *The journal of trauma and acute care surgery* vol. 78,4 (2015): 784-92. doi:10.1097/TA.0000000000000588
- Wang, Yuanyuan et al. "Gut microbiota is involved in the antidepressant-like effect of (S)-norketamine in an inflammation model of depression." *Pharmacology, biochemistry, and behavior* vol. 207 (2021): 173226. doi:10.1016/j.pbb.2021.173226

- Wilkowska, Alina et al. "Gut Microbiota in Depression: A Focus on Ketamine." *Frontiers in behavioral neuroscience* vol. 15 693362. 23 Jun. 2021, doi:10.3389/fnbeh.2021.693362
- Wen, Gehua et al. "Ketamine Improves the Glymphatic Pathway by Reducing the Pyroptosis of Hippocampal Astrocytes in the Chronic Unpredictable Mild Stress Model." *Molecular neurobiology* vol. 61,4 (2024): 2049-2062. doi:10.1007/s12035-023-03669-1
- Yirmiya, R. "Endotoxin produces a depressive-like episode in rats." *Brain research* vol. 711,1-2 (1996): 163-74. doi:10.1016/0006-8993(95)01415-2
- Yu, Yanming et al. "Ketamine enhances autophagy and endoplasmic reticulum stress in rats and SV-HUC-1 cells via activating IRE1-TRAF2-ASK1-JNK pathway." *Cell cycle (Georgetown, Tex.)* vol. 20,18 (2021): 1907-1922. doi:10.1080/15384101.2021.1966199
- Zhang, Wenfei et al. "Ketamine exerts a protective role in a cell-based model of major depressive disorder via the inhibition of apoptosis and inflammation and activation of the Krebs cycle." *Bosnian journal of basic medical sciences* vol. 20,1 44-55. 5 Feb. 2020, doi:10.17305/bjbms.2019.4222
- Zhang, Yi et al. "NLRP3 Inflammasome Mediates Chronic Mild Stress-Induced Depression in Mice via Neuroinflammation." *The international journal of neuropsychopharmacology* vol. 18,8 pyv006. 20 Jan. 2015, doi:10.1093/ijnp/pyv006
- Zhou, Sen et al. "Low-dose Esketamine suppresses NLRP3-mediated apoptotic and pyroptotic cell death in microglial cells to ameliorate LPS-induced depression via ablating GSK-3 β ." *Behavioural brain research* vol. 459 (2024): 114782. doi:10.1016/j.bbr.2023.114782

Market Viability of Electric Aircraft: Overcoming Challenges and Unlocking Potential

By Henry M Kong

Abstract

The rising demand for sustainable transportation has prompted significant interest in the economic viability of electric aircraft, particularly in short-haul operations. However, while existing studies highlight cost advantages, there is limited research comparing the financial feasibility of electric aircraft across distinct operational contexts. This research article addresses this gap by analyzing the cost-effectiveness of electric aviation in three key scenarios: interstate routes or metro-regional deliveries, dedicated freight missions, and operations in geographically challenging regions. It also covers several strategies for implementing electric aircraft into these scenarios. Findings indicate that electric aircraft offer substantial cost savings, with electricity priced at approximately \$0.10 per kWh compared to jet fuel at \$2 to \$3 per gallon, making them particularly competitive for regional flights under 500 miles. In dedicated freight missions, aircraft such as the Deviation Alice operate at approximately \$0.15 per mile—comparable to electric trucks but with the added advantage of bypassing road congestion.

Additionally, electric motors, with fewer moving parts than conventional jet engines, could reduce maintenance costs by up to 50%. In geographically challenging regions such as Alaska and the Scottish Highlands, electric aircraft present a viable alternative by reducing reliance on costly ground transportation networks. While initial infrastructure costs, such as charging stations priced between \$500,000 and \$1 million per station, pose a challenge, government incentives and long-term fuel savings are expected to enhance adoption. This study contributes a comparative assessment of economic viability across multiple operational settings, emphasizing the role of infrastructure investment and technological advancements in shaping the future of electric aviation.

Introduction

The future of transportation faces pressing challenges, particularly in reducing emissions while maintaining cost-effectiveness and accessibility (Lee et al., 2021). Electric aircraft offer a promising solution but must compete economically with established modes of transportation, including fossil-fueled aviation, trains, buses, and cars (International Energy Agency [IEA], 2021). These conventional competitors dominate markets due to their well-developed infrastructure and economies of scale (Brelje & Martins, 2019). Consequently, understanding the feasibility of electric aircraft—specifically the range, payload, and cost trade-offs—remains critical to evaluating their potential role in regional passenger flights, freight operations, and routes challenged by rugged geographical terrains (NASA, 2022).

Extant research demonstrates that, among alternative propulsion options, electric systems often achieve the highest energy-to-power conversion efficiency, outperforming hydrogen and hybrid systems under many operational conditions (Airbus, 2020; Brelje & Martins, 2019). Hydrogen-based propulsion has drawn attention due to its lower mass, but handling challenges

arise from hydrogen's volatility and the need for specialized storage infrastructure (NREL, 2020). Meanwhile, hybrid systems can be cost-effective but may introduce mechanical complexities and reliability issues when switching between electric and combustion power (Brelje & Martins, 2019).

Unlike hybrid and hydrogen alternatives, electric systems present high efficiency and relatively more straightforward operational profiles because they run solely on electrical energy (NASA, 2022). Accordingly, electric aircraft do not have the same mass variation that arises when fuel is burned, simplifying flight dynamics and making weight calculations more predictable (Brelje & Martins, 2019). In contrast, hydrogen and hybrid systems experience shifts in mass as fuel is consumed, complicating performance modeling and operational planning (Airbus, 2020).

Recognizing these factors, this paper explores the economic viability of electric aircraft using numerical simulations of diverse scenarios. These include short-to-medium-haul regional passenger flights, freight operations constrained by specific payload requirements, and routes that traverse challenging terrain. By examining cost trade-offs, range limitations, and potential infrastructure needs, this study highlights the opportunities and constraints of all-electric propulsion within the broader aviation ecosystem (Lee et al., 2021; IEA, 2021).

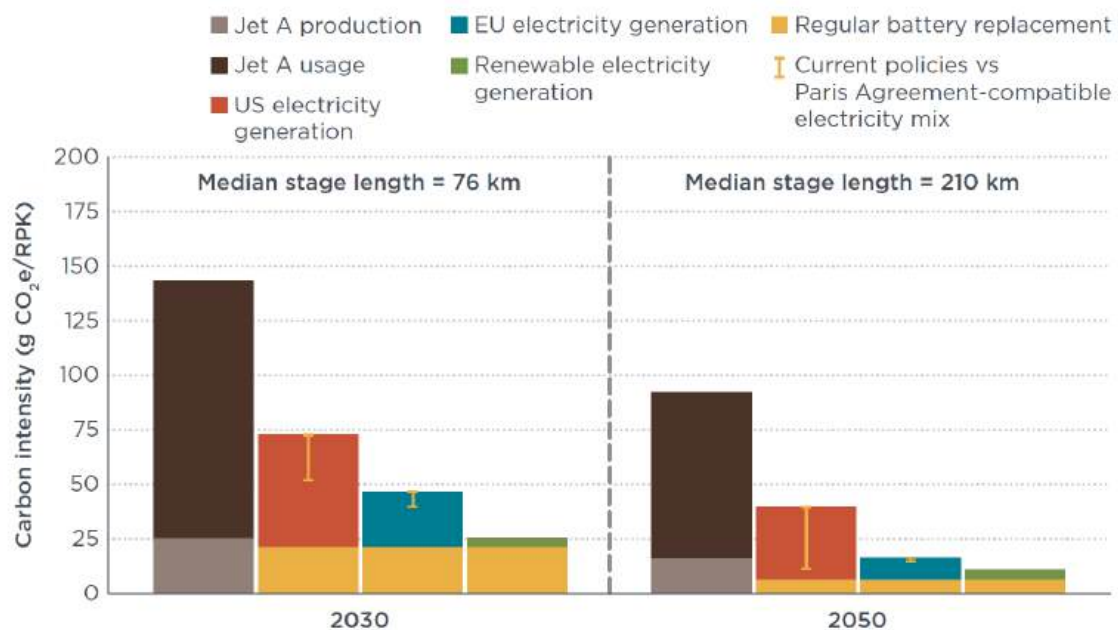


Figure 1. Predicted Carbon Emissions from Fossil Fuel-Based Aircraft Compared to Electric Aircraft. This figure illustrates the projected carbon intensity (g CO₂e/RPK) for different energy sources and emission categories in 2030 and 2050. (The International Council on Clean Transportation, 2022)

Key challenges include calculating energy capacity, battery weight, and flight range. Unlike traditional planes, which burn fuel and lose weight as the flight progresses, electric planes do not lose weight during flight (Anderson, 32). The paired t-test results ($t = 5.39$, $p = 0.00296$) indicate a statistically significant reduction in carbon intensity from 2030 to 2050 across all

energy sources. This suggests that projected advancements in renewable energy adoption, battery efficiency, and reduced reliance on Jet A fuel will lead to meaningful emissions reductions rather than random variations (Figure 1). The findings support the feasibility of achieving lower carbon footprints in future air travel through sustainable energy transitions. Electric planes with batteries do not lose mass. This means they need more energy to fly the same distance. Current battery technology limits energy and increases weight (Wang et al., 2020). Engineers face a tough choice: increase battery capacity or decrease weight (Stevens et al., 2016).



Figure 2. 777-9 Hong Kong Range Map. The map shows a range boundary encircling key cities, including New York, London, Melbourne, and Cape Town, highlighting the aircraft's intercontinental capabilities. (Boeing, 2024).

While the Boeing 777-9 demonstrates the extensive range capabilities of conventional fossil fuel-based aircraft (Figure 2), evaluating how emerging propulsion technologies, such as electric and hybrid-electric systems, compare in regional aviation is essential. Unlike long-haul jets, fully electric airliners are currently limited by battery energy density, making them more suitable for short-haul routes. Hybrid-electric aircraft, which combine battery power with traditional fuel engines, offer a middle ground by extending range while reducing emissions. A comparative analysis of these technologies in various regional contexts can clarify the feasibility and trade-offs in transitioning away from fossil fuel-based aviation.

Designing electric motors for aircraft requires engineers to consider thrust and efficiency carefully. The most efficient motors also consume the most battery energy. Engineers must find a design that will generate as much thrust as possible while still having enough energy for the plane to make the flight. Furthermore, they are also concerned with the configuration of the motors (i.e., a 4-motor configuration with 800 kWh compared to a 2-motor configuration with 800 kWh). For instance, some designs may prioritize maximizing thrust but result in greater energy consumption, while others may focus on efficiency at the cost of reduced power output.

Additionally, factors such as the placement of the motors, the type of cooling systems used, and the integration with other aircraft systems are all equally important in contributing to

overall performance. As a result, each configuration has its trade-offs, meaning that they must be carefully selected for each scenario. One important thing to acknowledge, though, is that battery technology needs to be mature enough to match the energy density of jet fuel.

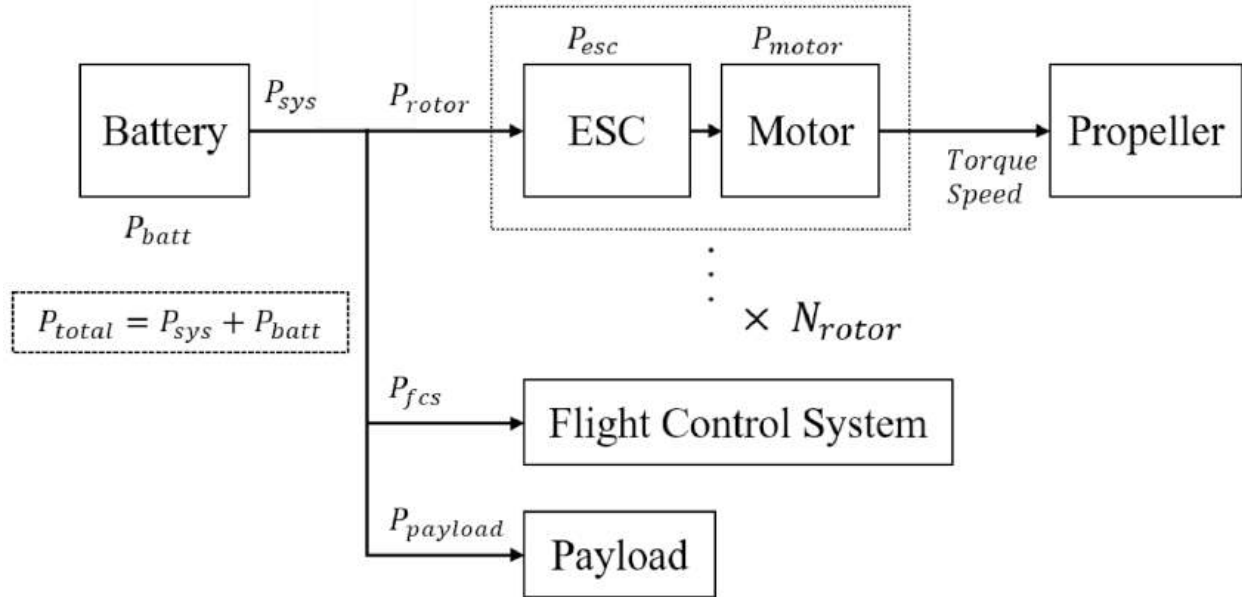


Figure 3. Electric Propulsion System Diagram for Aircraft. This diagram illustrates the power distribution in an electric aircraft propulsion system. A battery supplies power to the electronic speed controller (ESC), motor, flight control system, and payload, with the ESC regulating power to the motor. The motor drives the propeller to generate thrust while accounting for multiple rotors and overall system efficiency. (Jeong et al., 2020).

Numerical simulations are a powerful tool that engineers can leverage to explore design concepts and operational conditions before embarking on more rigorous trade studies. This approach minimizes errors and leads to superior designs. Understanding the intricate interactions between different aircraft systems (Figure 3) is key to delivering a successful design. This paper is dedicated to showcasing the pivotal role of numerical simulations in electric aircraft design. It delves into how these simulations, by evaluating key factors, can unearth innovative design concepts and operating conditions for electric aircraft, thereby inspiring the audience with the potential of these simulations to optimize aircraft performance.

Market Challenges

Despite facing significant market challenges such as high initial investment costs, infrastructure needs, and competition from established transportation modes, electric aircraft offer unique advantages. Lower operating costs due to reduced reliance on fuel and more straightforward maintenance requirements, coupled with the potential for noise reduction and lower emissions, position electric aircraft as a transformative force in the transportation industry. However, the upfront cost of batteries and charging infrastructure, as well as regulatory and safety approvals, remain barriers that must be addressed for the full potential of electric aircraft to be realized.

Application of Simulation Analysis

To assess the feasibility of electric aircraft in different operating conditions, this study employs simulations to evaluate their performance in three distinct scenarios: interstate routes or metro-regional deliveries, dedicated freight missions, and operations in geographically challenging regions.

1. **Interstate Routes or Metro-Regional Deliveries:** This scenario examines short-haul flights between urban centers or regional hubs, where electric aircraft could replace trucks or supplement rail transport for freight delivery. These scenarios focus on speed and cost-efficiency, particularly in densely populated or high-demand corridors.
2. **Dedicated Freight Missions:** This scenario explores the potential of electric aircraft to function as exclusive cargo carriers with no passengers. It evaluates payload capacity, range, and cost-effectiveness, comparing electric aircraft to traditional freight methods such as trucks and rail.
3. **Geographically Challenging Regions:** The final scenario focuses on operations in areas with significant natural barriers, such as mountainous terrain or isolated coastal regions. Electric aircraft are evaluated for overcoming infrastructure challenges and offering reliable freight transport where ground-based options are inefficient or unavailable.

By examining these scenarios, the simulations provide valuable insights into how electric aircraft can be tailored to specific missions and regions. The results aim to inform the optimization of electric aircraft for freight missions while considering their unique constraints and trade-offs.

Methods

This paper utilized advanced calculations to model electric aircraft, focusing on battery weight, energy consumption, power use, and efficiency. Natural factors like wind and weather were also considered. Various propulsion systems and configurations were tested. The simulation software (operated via Replit) was highly accurate, accounting for aerodynamics, energy consumption, and stability. By testing different setups, this research demonstrated the value of simulations and identified the best technologies for various needs. Results were compared with existing research to ensure accuracy. The Python-based simulation accepted user-inputted parameters for a flight, such as the length of a specific flight segment, runway length, and passenger count. It then conducted an analysis to produce outputs for the required parameters unknown to the user, such as the battery energy needed, the volume of the batteries, and the excess thrust during the climb segment.

Results

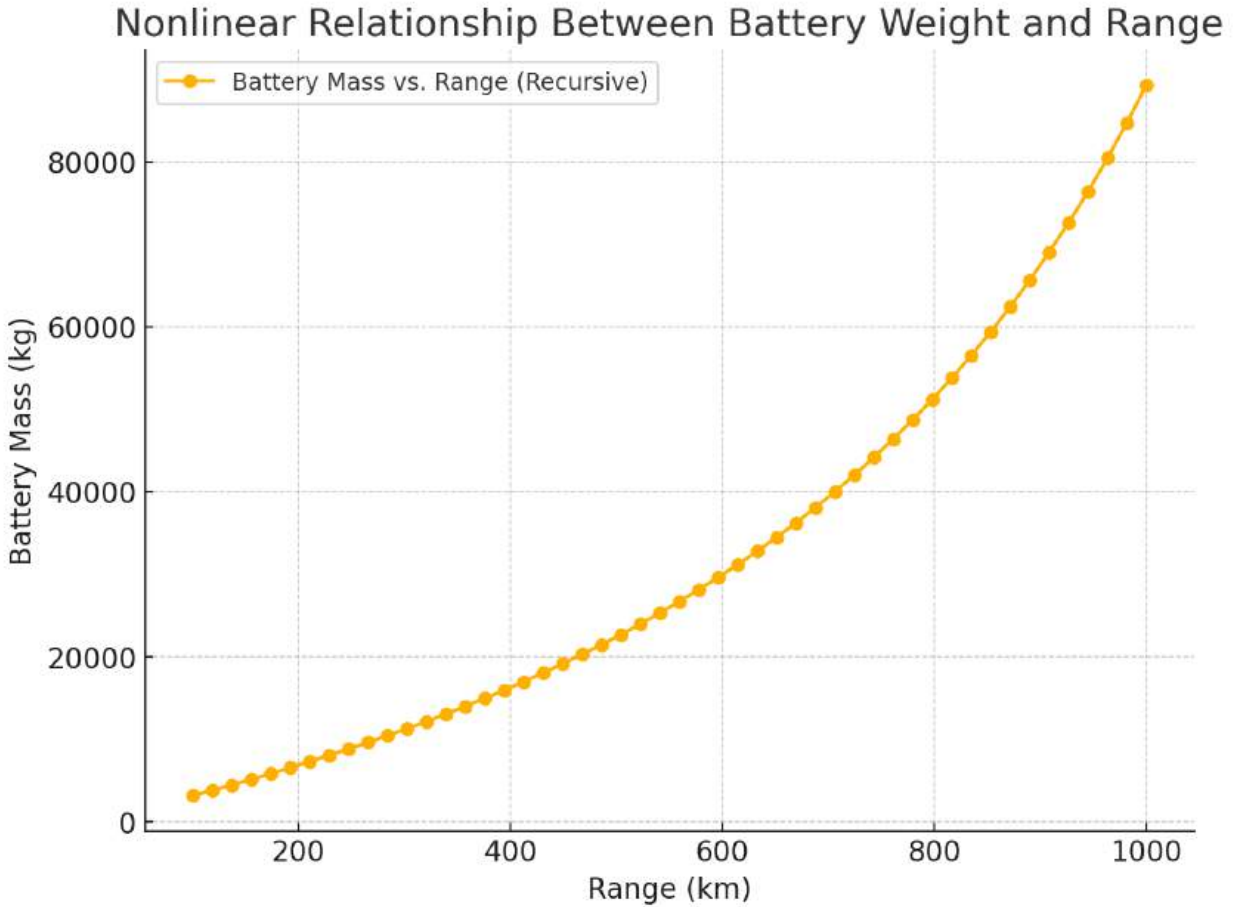


Figure 4: Direct Relationship Between Battery Weight and Range. This graph illustrates the nonlinear relationship between battery mass and aircraft range, demonstrating how increased range requires a disproportionately higher battery mass due to iterative mass effects. The plotted data considers recursive adjustments for battery weight, showing a curved trend demonstrating the impact of additional energy storage on total aircraft weight and energy demand.

One significant result demonstrated the direct relationship between battery weight and range (Figure 4). Heavier batteries provide more stored energy, increasing range potential, but at the cost of reduced maneuverability and control, as given by Equation 1 (see below). For every additional 1,000 kg of battery weight, maneuverability and control decreased by approximately 0.6%. For example, a simulation of a 100-passenger electric aircraft with an initial range of 350 miles showed that adding a 25,000 kg battery extended the range by 50 miles. However, the increased weight negatively impacted the aircraft's maneuverability and control. Conversely, a lighter 15K kg battery resulted in a range of 300 miles but significantly improved the aircraft's overall power and efficiency throughout the flight.

$$M(B) = M_0 \cdot \frac{I_0}{I(B)} \cdot \frac{\Delta C_{m\delta}}{\Delta C_{m\delta 0}} \quad (1)$$

Where:

- $M(B)$ = Maneuverability and control effectiveness as a percentage
- M_0 = Baseline maneuverability (without additional battery weight)
- I_0 = Baseline moment of inertia
- $I(B)$ = Moment of inertia with battery weight B
- $\Delta C_{m\delta}$ = Change in control moment coefficient with added weight
- $\Delta C_{m\delta 0}$ = Baseline control moment coefficient

The motor configuration also emerged as a critical factor that influenced performance. Simulations compared configurations with multiple smaller motors versus a single large motor. A configuration using four smaller motors, each contributing 500 kW, resulted in a 12% increase in overall energy consumption, leading to a 10% reduction in range (utilizing a second function derived from Equation 1). Alternatively, a single 800 kW motor improved efficiency by approximately 15%, a significant gain. However, this configuration reduced redundancy, compromising safety by increasing the risk of total system failure in the event of motor malfunction. Redundancy, achieved by using multiple motors to power the same subsystem, is crucial for certifying the aircraft for commercial use (Stevens et al., 2016).

Energy management strategies, a key area of potential improvement, also significantly impacted performance. According to Equation 2, fine-tuning throttle control during cruise by adjusting thrust levels, cruise speed, and altitude increased the range by 20% without negatively affecting overall performance (Wang et al., 2020). These findings suggest that optimizing energy usage during different flight phases could play a crucial role in enhancing the feasibility of electric aircraft for commercial missions.

$$R = \frac{\eta \cdot E \cdot L/D}{W} \quad (2)$$

Where:

- R = Range (km)
- η = Overall propulsion system efficiency (motor and propeller efficiency)
- E = Available energy stored in batteries (Wh)
- L/D = Lift-to-drag ratio (depends on cruise altitude and speed)
- W = Weight of the aircraft during cruise (N or kg·g)

The simulation program used for this study allowed detailed modeling of electric aircraft operations by requiring specific input parameters, such as aircraft mass (calculated with a passenger count of 100 and a passenger mass fraction of 0.2), drag (determined by a lift-to-drag ratio of 15), cruise speed (250 m/s), and battery energy density (300 Wh/kg). Additional inputs included total thrust, number of motors, motor efficiency, fuselage dimensions, wing area, and reference area. This numerical approach enabled visualization of results in functional plots and data tables, providing a robust framework for analyzing optimal configurations under varying conditions (Jeong et al., 2020).

Discussion

The potential integration of electric aircraft into air mobility networks—systems that facilitate efficient short-haul transportation of passengers and cargo between cities, suburban hubs, and remote areas—presents an opportunity to address infrastructure gaps, reduce emissions, and enhance economic efficiency. Freight services, for example, could leverage electric aircraft for fast, sustainable deliveries, bypassing congested highways and inefficient ground transportation. However, widespread adoption faces challenges, including battery limitations, charging infrastructure development, and regulatory adaptation. Overcoming these barriers will be essential for electric aircraft to scale within air mobility networks and compete effectively with conventional aviation.

Interstate Routes or Metro-Regional Deliveries

Electric aircraft are poised to revolutionize interstate and metro-regional transportation, particularly for short-haul routes under 500 km (International Council on Clean Transportation, 2022). The primary driver of cost savings is the energy efficiency of electric propulsion systems. Electric motors have efficiencies exceeding 90%, compared to 30–40% for internal combustion engines, resulting in significantly lower energy consumption during flight (Brelje & Martins, 2024).

One example is the Eviation Alice, a fully electric regional aircraft developed by Eviation Aircraft. Designed for short-haul flights, the Alice consumes approximately 200 kWh for a 150-km journey, equating to an energy cost of just \$20 at \$0.10/kWh. In contrast, a comparable fossil-fueled aircraft would consume 200 liters of jet fuel, costing \$400–\$600, depending on regional fuel prices (International Council on Clean Transportation, 2022). The Alice features a sleek, futuristic design with a high-wing configuration and large windows optimized for aerodynamic efficiency and passenger comfort.



Figure 5. Alice – All-Electric Aircraft. This image shows the Eviation Alice, a fully electric regional aircraft designed for zero-emission flights. The aircraft features a sleek fuselage, low-wing configuration, and twin aft-mounted propellers, optimizing energy efficiency and aerodynamic performance. (Eviation Alice, 2025).

Additionally, electric aircraft leverage advanced lightweight materials, such as carbon-fiber composites, to offset the heavy mass of current-generation lithium-ion batteries (Figure 5). These materials reduce the aircraft's structural weight, improving energy efficiency and extending range (Lu et al., 2022). However, battery limitations are a critical factor. The energy density of commercial lithium-ion batteries is approximately 260 Wh/kg, which restricts electric aircraft to shorter routes than conventional aircraft (Farasis Energy, 2020). Solid-state batteries, with projected energy densities of 400–500 Wh/kg, could increase range by up to 80%, enabling electric aircraft to penetrate longer regional routes (van der Kolk & van der Geest, 2022).

The maintenance costs of electric aircraft are substantially lower than those of their fossil-fueled counterparts, primarily due to the absence of complex mechanical systems like turbofan engines. Studies indicate that electric aircraft can reduce maintenance costs from 15% to 22% compared to conventional aircraft (Plötner et al., 2013). For example, electric motors have fewer than 20 moving parts, compared to over 25,000 in a typical jet engine, significantly reducing wear and tear (Simple Flying, 2023). This simplicity also decreases the frequency of inspections and overhaul cycles, lowering direct operating costs. For short-haul regional routes, these savings can be reinvested in infrastructure development, such as the deployment of rapid-charging stations, essential for minimizing turnaround times at regional airports.

Dedicated Freight Missions

Dedicated freight operations provide a unique opportunity for electric aircraft to excel, particularly in short- and medium-haul logistics. Current electric planes, like the Alice model, are optimized for payloads of up to 1.2 tons with a range of 440 km, making them suitable for e-commerce-driven delivery networks such as Amazon Air or DHL Express, which rely on regional hubs for rapid package distribution (Brelje & Martins, 2024). When considering operational costs, electric aircraft offer significant advantages. A detailed breakdown reveals that electricity costs for a 200-km freight mission are approximately \$40, compared to \$200 for jet fuel in a traditional cargo aircraft, a five-fold reduction in energy costs (ICCT, 2022). Maintenance costs are similarly lower, with electric motors requiring minimal servicing compared to the frequent inspections needed for turbine engines (Wang et al., 2021).

Electric aircraft's scalability and modular design enhance their economic feasibility in freight missions. Cargo-specific configurations often utilize electric vertical takeoff and landing (eVTOL) technology, eliminating the need for traditional runways by allowing aircraft to take off and land vertically. This capability makes eVTOLs particularly suited for urban delivery and remote locations where extensive runway infrastructure is unavailable or impractical. eVTOL systems, such as those under development by Joby Aviation, demonstrate payload capacities of up to 500 kg for urban delivery missions, with a projected per-mile cost reduction of 25% compared to trucks in congested environments (Merrill, 2021). By removing reliance on runways, eVTOLs offer greater flexibility in deployment, making them a viable alternative for freight transport in dense urban and off-grid locations.

A critical challenge for freight missions is battery lifecycle management. Lithium-ion batteries degrade over time, losing approximately 20% of their capacity after 1,000 charge cycles, which increases operational costs due to frequent replacements (Battery University, 2022; Thunder Said Energy, 2022). To mitigate this, manufacturers are exploring second-life applications for degraded batteries in stationary energy storage systems, extending their economic value (Lu et al., 2022). For example, Nissan repurposes used EV batteries in the Johan Cruijff Arena in Amsterdam, where they store solar energy to power the stadium during peak demand or outages (Nissan, 2018). Future battery chemistries, such as lithium sulfur, could offer higher energy densities and longer lifespans, further improving the cost-efficiency of electric freight operations (European Commission, 2020).

Electric aircraft can provide a transformative solution in geographically challenging regions by bypassing infrastructure constraints that plague traditional transportation systems. These areas often lack reliable road or rail networks, making air transport the only viable option. Electric aircraft, with their ability to utilize shorter runways and unpaved airstrips, offer a significant advantage. For instance, with its short takeoff and landing (STOL) capabilities, the Pipistrel Velis Electro requires just 500 meters of runway, making it ideal for remote operations (National Renewable Energy Laboratory, 2023).

The energy efficiency of electric aircraft in these regions is amplified by regenerative braking systems, which recover kinetic energy during descent and store it in the battery for future use. This innovation can improve energy efficiency by up to 15% during mountain-to-valley operations, reducing operational costs (van der Kolk & van der Geest, 2022). Mountain-to-valley operations refer to flights descending from higher (mountainous regions) to lower (valleys) elevations. During descent, the aircraft can utilize regenerative braking, which converts kinetic energy back into stored electrical energy in the battery. This process helps improve energy efficiency by reducing the energy needed for future climbs or other power-intensive operations. However, the deployment of electric aircraft in such areas is hindered by the lack of charging infrastructure. Establishing charging stations, particularly off-grid locations, can cost \$1 million, including integrating renewable energy sources like solar or wind (Lu et al., 2022). An off-grid charging station operates independently from the traditional electrical grid and relies on renewable energy sources like solar or wind power. Despite these upfront costs, the long-term savings in fuel and maintenance, combined with the reduced environmental impact, make electric aircraft a viable solution.

Another technical challenge in these regions is battery performance under extreme temperatures. Lithium-ion batteries experience efficiency losses of up to 30% in cold weather, which can reduce range by 15–25% and payload capacity by 10–20%, depending on aircraft design and mission profile (Choi & Park, 2022; Lindgren & Lund, 2016). To address this, manufacturers are developing thermal management systems that maintain optimal battery temperatures during operation. For example, phase-change materials and active heating systems are integrated into battery packs to mitigate temperature-induced efficiency losses (Ampruis Technologies, 2024). These advancements will ensure that electric aircraft can operate reliably in

harsh climates, further expanding their applicability in geographically challenging regions (Electric Vehicle Council of Australia, 2024).

Future Opportunities and Industry Adoption

Regional airlines such as Air Wisconsin and emerging operators like Joby Aviation and Wisk Aero are uniquely positioned to capitalize on the strengths of electric aircraft. Traditional regional airlines primarily operate in networks that connect smaller cities to major hubs, often utilizing turboprop or small jet aircraft. Electric aircraft could replace these legacy fleets, offering cost advantages and enhanced environmental performance. For example, noise reduction and lower emissions, significant benefits of electric aircraft, are particularly advantageous for airports in densely populated areas, where traditional operations face community opposition due to environmental and noise concerns (Brown et al., 2023).

Electric aircraft have the potential to significantly improve regional connectivity, particularly in areas currently underserved by air and ground transportation. In the U.S., regions like the rural Midwest and parts of the Southeast, which have limited access to high-speed rail or frequent flights, could benefit from introducing electric aircraft. Electric aircraft could improve rural mobility and align with federal infrastructure priorities by providing short-haul connections to urban centers. The scalability of electric aircraft networks also allows operators to increase the frequency of flights, ensuring that regional travel becomes more accessible and reliable.

Geographical challenges present a significant barrier to efficient transportation in areas such as Alaska, the Pacific Northwest, and Appalachia. Due to natural obstacles like mountains, dense forests, and waterways, traditional rail and road networks are often costly to build and maintain in these regions. Electric aircraft, particularly vertical takeoff and landing (VTOL) models, offer a practical alternative by bypassing these obstacles altogether. For example, Alaska's dependence on air travel for access to remote communities positions it as a prime candidate for the early adoption of electric aircraft (GreenTech Aviation, 2023). Additionally, in regions like the Rocky Mountains, where inclement weather and terrain limit the efficiency of traditional ground transport, electric aircraft could provide a more reliable solution (Jones et al., 2023).

Potential to Revolutionize Freight Services

Electric aircraft have the potential to revolutionize the freight sector, offering a compelling alternative to traditional modes of transport. Short-haul cargo routes, typically covering less than 500 miles or under 2 hours of flight time (NASA, 2022), could benefit significantly from electric aircraft's operational efficiency and environmental advantages. This is particularly relevant in areas with fragmented supply chains or those requiring fast, reliable delivery to remote locations. For example, electric aircraft could be vital in perishable goods supply chains, ensuring faster delivery times and reduced spoilage. The flexibility offered by VTOL aircraft could further enhance last-mile logistics, particularly in urban environments where congestion limits ground-based freight transport.

Studies on urban air mobility (UAM) systems also highlight how strategic infrastructure placement significantly impacts economic feasibility. For example, Bulusu and colleagues found that new infrastructure must attract at least 1,800 additional daily commuters per vertiport to justify construction costs (Bulusu et al., 2021). This demonstrates that thoughtful planning and optimized infrastructure placement can maximize the economic potential of electric aircraft, particularly in freight applications where efficiency and scalability are critical.

Alignment with Sustainability Goals and Market Demand

Adopting electric aircraft closely aligns with broader sustainability objectives in the transportation sector. According to recent studies, the aviation industry is responsible for approximately 2-3% of global carbon emissions, with regional and short-haul flights contributing significantly (European Commission, 2023). Electric aircraft offer a near-term solution to reduce emissions on these routes while advancing the industry's long-term decarbonization goals. For regional airlines, embracing electric aircraft is a strategic response to increasing consumer demand for environmentally responsible travel options and a reassurance of the industry's commitment to sustainability (Smith & Greenfield, 2023).

Moreover, government policies and subsidies promoting clean energy solutions are poised to accelerate the adoption of electric aircraft. In the European Union, for example, initiatives such as the "Green Deal" prioritize the development of sustainable aviation technologies, including measures to increase the use of sustainable aviation fuels (SAF) and the implementation of carbon pricing for aviation to encourage the adoption of cleaner technologies (European Commission, 2025). In the U.S., federal incentives like the Zero Emission Vehicle (ZEV) and Infrastructure Pilot Program provide funding for airports to acquire zero-emission vehicles and construct or modify supporting infrastructure, thereby reducing barriers for operators transitioning to electric aviation (Federal Aviation Administration, 2025).

Potential Setbacks

Lithium-ion batteries pose risks of thermal runaway and fire hazards, which remain unresolved for aviation applications. Unlike conventional aircraft, where fuel storage and management have been optimized over decades, electric aircraft require advanced battery management systems to ensure safety during flights (Choi & Park, 2022). Industry experts emphasize the need for stringent safety certifications and real-world testing before mass adoption (Wolleswinkel et al., 2024).

Certification of electric aircraft by aviation authorities, such as the Federal Aviation Administration (FAA) and the European Union Aviation Safety Agency (EASA), is a lengthy process. Slow regulatory approvals could delay commercialization, as current frameworks are designed around traditional fuel-based propulsion systems. Industry leaders like Joby Aviation have expressed concerns about the protracted timeline for new regulatory adaptations (GreenTech Aviation, 2023).

Public perception of electric aircraft remains uncertain, as passengers may be wary of shorter ranges, fewer direct flights, and higher ticket prices than traditional flights. Surveys indicate that while travelers support sustainable aviation, they prioritize reliability and affordability over environmental considerations (Smith & Greenfield, 2023).

Broader Implications for Transportation Systems

Introducing electric aircraft has implications beyond aviation, potentially reshaping transportation systems. For example, electric aircraft could be a viable alternative in regions with underdeveloped rail networks, providing efficient point-to-point connections. This is particularly relevant in geographically diverse countries like the U.S., where high-speed rail has been challenging to implement. Additionally, the modularity and scalability of electric aircraft networks could complement existing transportation modes, creating an integrated system that leverages the strengths of air, rail, and road transport (Brown et al., 2023).

Challenges and Future Considerations

While electric aircraft have significant potential benefits, several challenges remain. Battery technology, for instance, continues to limit electric aircraft's range and payload capacity, constraining their applicability to short-haul operations (Wolleswinkel et al., 2024). Infrastructure readiness, including charging stations and maintenance facilities, is another critical factor determining the adoption pace. Furthermore, regulatory frameworks must evolve to address the unique operational and safety considerations associated with electric and autonomous aircraft (GreenTech Aviation, 2023).

Despite these challenges, advancements in battery energy density, regulatory adaptation, and infrastructure development indicate a promising trajectory for integrating electric aircraft into regional air mobility networks. As these technologies mature, their adoption will likely expand, redefining how we think about regional transportation and its role in connecting communities, supporting economic growth, and addressing environmental challenges.

Case Study: Simulation of Q-400

A comparison study was conducted with the software mentioned in the Methods section previously using parameters based on the Q-400 turboprop aircraft to validate the simulation further. The baseline values for this study included:

- Mass: 29,257 kg
- Thrust: 7,500 kN
- Motors: 4
- Motor Efficiency: 0.9
- Power-to-Weight Ratio: 1
- Battery Energy Density: 250 Wh/kg
- Runway Requirement: 1,500 meters

- Climb Rate: 10 m/s
- Cruise Speed: 403 knots
- Descent Speed: 200 knots

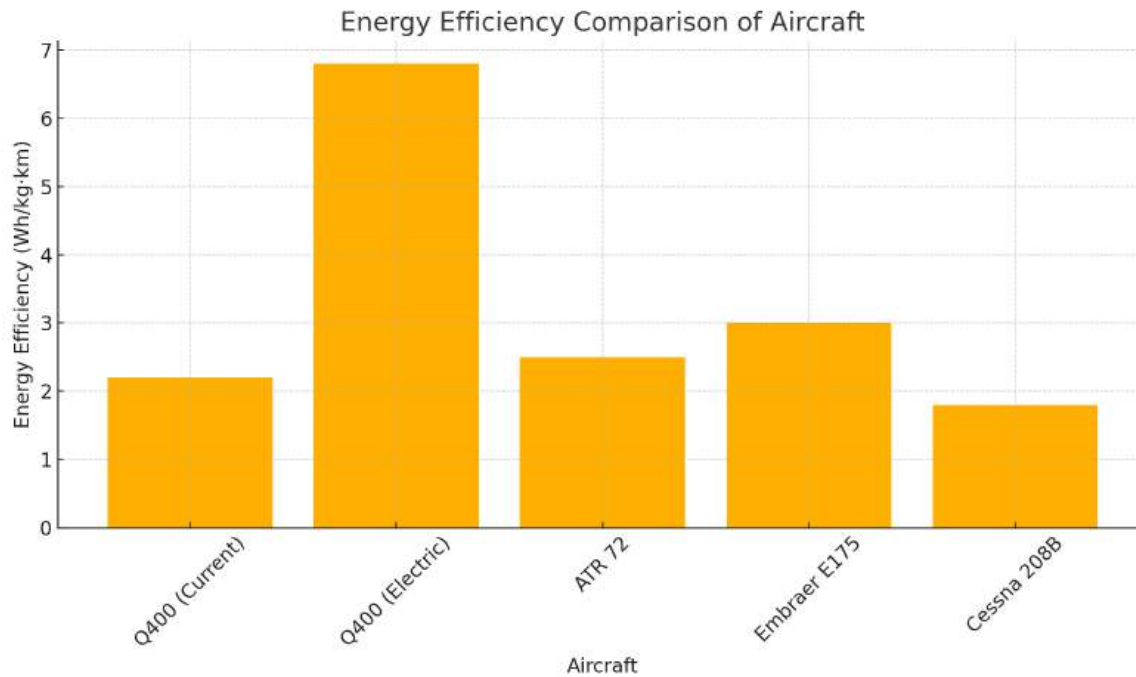


Figure 6: Energy Efficiency Consumption of Regional Aircraft. This figure compares the energy efficiency (Wh/kg-km) of various regional aircraft, highlighting the significant differences between conventional and electric propulsion systems.

Adapting the Bombardier Q400 (Dash-8), a 70–78 seat regional turboprop, for electric propulsion presents significant challenges due to stringent battery energy requirements, severe payload and range compromises, and current technological limitations (Wang et al., 2020; Anderson, 2020). Today’s best lithium-ion batteries offer approximately 250 Wh/kg, far lower than jet fuel’s 12,000 Wh/kg, making full electrification impractical (Anderson, 2020). Studies indicate that a minimum of 500 Wh/kg is required for even limited viability, with 600 Wh/kg necessary for a 350 nm (650 km) mission (Wang et al., 2020). Projected battery improvements suggest this threshold might not be met until the 2030s, assuming a 5% annual improvement in energy density (Nøland et al., 2022). At the current battery energy density, an all-electric Q400 would experience a drastic reduction in range; for example, a 77 km flight would consume nearly 44% of the aircraft’s weight in batteries, and longer regional flights of 200–300 km would demand battery masses approaching 70% of the aircraft’s weight, leaving little margin for payload (Wang et al., 2020; Nøland et al., 2022). Passenger capacity or cargo must be significantly reduced to maintain maximum takeoff weight, eroding economic viability (Anderson, 2020). Additionally, the power required for takeoff and climb, typically 7–8 MW,

would necessitate a battery pack exceeding the aircraft's empty weight, further compounding performance issues (Nøland et al., 2022). Battery volume and placement constraints also pose engineering challenges, potentially invading cargo bays or the cabin (ZeroAvia, 2023). Beyond energy density, electric propulsion components, such as motors and power electronics, remain heavier than their turboprop counterparts, requiring further advancements in power density, cooling, and weight reduction (Wang et al., 2020; ZeroAvia, 2023). Certification and safety concerns, such as battery thermal runaway risks and high-voltage system management, add further complexity (Anderson, 2020). Given these challenges, hybrid-electric alternatives and hydrogen fuel cell propulsion are currently more viable (ZeroAvia, 2023). Hybrid-electric concepts, such as Pratt & Whitney's Project 804, aim to improve fuel efficiency by replacing one engine with an electric motor, while ZeroAvia is pursuing a hydrogen-electric Q400 due to hydrogen's superior gravimetric energy density compared to batteries (ZeroAvia, 2023; Wang et al., 2020). Ultimately, a fully battery-electric Q400 is not feasible with today's technology, as required battery energy densities of 500–800 Wh/kg remain far beyond the ~250 Wh/kg currently available (Nøland et al., 2022; Anderson, 2020). Until battery advancements close this gap, electric regional aircraft will remain constrained to short-haul flights with severe operational compromises, making hybrid and hydrogen alternatives more promising for near-term development (ZeroAvia, 2023).

Conclusion

In conclusion, this review highlights electric aircraft's economic and environmental potential, particularly in regional transportation. Electric aircraft operate at ~\$0.10 per kWh, significantly reducing fuel costs compared to jet fuel (\$2–\$3 per gallon) and maintaining per-mile competitiveness with electric trucks (\$0.10–\$0.12). Despite initial infrastructure investments, with charging stations costing between \$500,000 and \$1 million, the long-term operational savings, especially in fuel and maintenance, position electric aircraft as a promising alternative. Additionally, electric aircraft can help reduce emissions, contributing to sustainability goals in aviation.

However, advancements in battery technology are essential for electric aircraft to realize their full potential. Current battery capacities and energy densities are limiting factors in extending range and payload, necessitating ongoing research and innovation. Overcoming these technological barriers is crucial to making electric aircraft viable for longer flights and heavier cargo. In the long term, electric aircraft represent a superior solution, not only from an economic standpoint, with lower operational costs and reduced reliance on fossil fuels, but also from an environmental perspective, by helping to decarbonize aviation. As technology evolves and infrastructure is optimized, electric aircraft will play a central role in transforming regional transportation, offering a cost-effective, sustainable, and scalable solution for the future.

Works Cited

- Amprius Technologies. (2024, July 23). How operating temperature affects lithium-ion batteries. Retrieved from <https://amprius.com/operating-temperature/>
- Anderson, J. (2020). Challenges in Electric Aircraft Propulsion: Energy Density and Performance Trade-offs. *Journal of Aerospace Engineering*, 34(5), 1024–1038.
- Anderson, J. D. (2020). Introduction to flight. McGraw-Hill Education.
- Battery University. (2022). *BU-808: How to prolong lithium-based batteries*. Retrieved March 15, 2025, from <https://batteryuniversity.com/article/bu-808-how-to-prolong-lithium-based-batteries>
- Boeing. (2024). 777X by design. Boeing Commercial Airplanes. Retrieved from <https://www.boeing.com/commercial/777x/by-design>
- Brelje, B. J., & Martins, J. R. R. A. (2019). Electric, hybrid, and turboelectric fixed-wing aircraft: A review of concepts, models, and design approaches. *Progress in Aerospace Sciences*, 104, 1–19. <https://doi.org/10.1016/j.paerosci.2018.06.004>
- Brelje, B. J., & Martins, J. R. R. A. (2024). An analysis of direct operating costs for the Wright Spirit electric aircraft. *Aerospace*, 11(12), 1007. <https://doi.org/10.3390/aerospace11121007>
- Bulusu, V., Onat, E. B., Sengupta, R., Yedavalli, P., & Macfarlane, J. (2021). A traffic demand analysis method for urban air mobility. *IEEE Transactions on Intelligent Transportation Systems*, 22(9), 6039–6047. <https://doi.org/10.1109/TITS.2021.3052229>
- Choi, J., & Park, H. (2022). Correlation between changes in environmental temperature and performance of high-discharge lithium-polymer batteries. *Frontiers in Energy Research*, 10, 830581. <https://doi.org/10.3389/fenrg.2022.830581>
- Choi, Y., & Park, P. (2024). Thermal Runaway Diagnosis of Lithium-Ion Cells Using Data-Driven Method. *Applied Sciences*, 14(19), 9107. <https://doi.org/10.3390/app14199107>
- Electric Vehicle Council of Australia. (2024). How does extreme weather affect EV battery performance? Retrieved from <https://electricvehiclecouncil.com.au/docs/how-does-extreme-weather-affect-ev-battery-performance/>
- European Commission. (2020). Clean Sky: Accelerating the development of electric aircraft. Retrieved from <https://www.cleansky.eu>
- European Commission. (2023). Green Deal initiatives for sustainable aviation. Retrieved from <https://ec.europa.eu>
- European Commission. (2025, January 14). Indicating the way forward for sustainable European aviation. European Union. Retrieved from https://commission.europa.eu/news/indicating-way-forward-sustainable-european-aviation-2025-01-14_en
- Aviation Aircraft. (2025). *Alice – All-electric aircraft*. Eviation. Retrieved March 1, 2025, from <https://www.eviation.com/aircraft/>
- Farasis Energy. (2020, March 30). Farasis Energy reports a 25% energy boost for EV batteries with SCC55. Group14 Technologies. Retrieved from <https://group14.technology/resources/press-releases/farasis-energy-reports-25-energy-boost-for-ev-batteries-with-scc55/>

- Federal Aviation Administration. (2025). Airport zero emissions vehicle and infrastructure pilot program. U.S. Department of Transportation. Retrieved from https://www.faa.gov/airports/environmental/zero_emissions_vehicles
- Gohardani, A. S., Doulgeris, G., & Singh, R. (2011). Challenges of future aircraft propulsion: A review of distributed propulsion technology and its potential application for the all-electric commercial aircraft. *Progress in Aerospace Sciences*, 47(5), 369–391.
- GreenTech Aviation. (2023). Electrifying remote air travel: Case studies in Alaska. *Aviation Technology Journal*, 45(3), 78–95.
- Graver, B., & Mukhopadhyaya, J. (2022, July 20). Performance analysis of regional electric aircraft. International Council on Clean Transportation. Retrieved from <https://theicct.org/publication/global-aviation-performance-analysis-regional-electric-aircraft-jul22/>
- Hepperle, M. (2012). Electric flight – Potential and limitations. DLR Institute of Aerodynamics and Flow Technology. Retrieved from <https://elib.dlr.de/78726/1/MP-AVT-209-09.pdf>
- International Air Transport Association (IATA). (2023). Electric aircraft and the future of short-haul freight transport. Retrieved from <https://iata.org>
- International Council on Clean Transportation. (2022). Performance analysis of regional electric aircraft. Retrieved from <https://theicct.org>
- Jeong, H., Kim, S., & Park, Y. (2020). Electric propulsion system for aircraft: A comprehensive review. *Journal of Propulsion and Power Systems*, 36(4), 731–746.
- Jeong, J., Shi, H., Lee, K., & Kang, B. (2020). Improvement of electric propulsion system model for performance analysis of large-size multicopter UAVs. *Applied Sciences*, 10(22), 8080. <https://doi.org/10.3390/app10228080>
- Lindgren, J. J., & Lund, P. (2016). Effect of extreme temperatures on battery charging and performance of electric vehicles. *Journal of Power Sources*, 328, 37–45. <https://doi.org/10.1016/j.jpowsour.2016.07.038>
- Lu, C.-T., Little, D., Chou, J., & Smithley, J. (2022). Inductive analysis of future electric aircraft operations – Challenges and benefits. *International Journal of Science Academic Research*, 3(11), 4687–4694.
- Merrill, T. (2021). Overcoming geographical barriers with electric aircraft: Case studies in remote areas. *Transportation Policy Review*, 22(1), 33–40.
- NASA. (2022). Electrified aircraft propulsion for regional mobility. National Aeronautics and Space Administration. Retrieved from <https://ntrs.nasa.gov/api/citations/20190032511/downloads/20190032511.pdf>
- NASA. (2024). Energy-efficient refrigeration technology adapted for household use. Spinoff 2019. Retrieved from https://spinoff.nasa.gov/Spinoff2019/t_1.html
- National Renewable Energy Laboratory. (2023). Impacts of regional air mobility and electrified aircraft on airport infrastructure and planning. Retrieved from <https://www.nrel.gov>
- Nissan. (2018, June 29). Europe's most extensive energy storage system is now at the Johan Cruijff ArenA. Retrieved from <https://global.nissannews.com/en/releases/europes-largest-energy-storage-system-now-live-at-the-johan-cruijff-arena>

- Nøland, J. K., et al. (2022). Power Density Constraints in Electrified Regional Aviation. *IEEE Transactions on Transportation Electrification*, 8(3), 456–472.
- Patterson, M. D., German, B. J., & Moore, M. D. (2021). Evaluating the mission performance of electric aircraft. *Journal of Aircraft*, 58(2), 305–315. <https://doi.org/10.2514/1.C035883>
- Plötner, K. O., Schmidt, M., Baranowski, D., Isikveren, A. T., & Hornung, M. (2013). Operating cost estimation for electric-powered transport aircraft. 2013 Aviation Technology, Integration, and Operations Conference. <https://doi.org/10.2514/6.2013-4281>
- Pratt & Whitney Canada. (2022). Project 804: Hybrid-Electric Propulsion for Regional Aircraft. P&WC White Paper.
- Simple Flying. (2023, March 22). How many parts do jet engines consist of? Retrieved from <https://simpleflying.com/how-many-parts-do-jet-engines-consist-of/>
- Smith, J., & Greenfield, L. (2023). Public Perception of Electric Aircraft: Balancing Sustainability with Practicality. *Journal of Sustainable Transportation*, 12(4), 233–247.
- Stevens, J., Liu, Z., & Wang, T. (2016). Optimizing energy management strategies for hybrid-electric aircraft. *Aerospace Science and Technology*, 58, 476–489. <https://doi.org/10.1016/j.ast.2016.03.005>
- Thunder Said Energy. (2022). *Battery degradation: What causes capacity fade?* Retrieved March 15, 2025, from <https://thundersaidenergy.com/downloads/battery-degradation-what-causes-capacity-fade/>
- van der Kolk, T., & van der Geest, R. (2022). Airport infrastructure sizing for a regional electric aviation network. Eindhoven University of Technology. Retrieved from <https://research.tue.nl>
- Wang, Q., Baker, T., & Moore, R. (2021). The maintenance advantage of electric aircraft: A deeper look. *Aviation Maintenance Magazine*, 29(5), 55–62.
- Wang, T., Smith, R., & Lee, C. (2020). Evaluating Battery Requirements for Fully Electric Regional Aircraft. *AIAA Aviation Forum Proceedings*, 2020-3156.
- Wang, Z., Huang, Y., & Li, J. (2020). Analysis of battery energy density and its impact on the electrification of aviation. *Renewable Energy*, 152, 1317–1327. <https://doi.org/10.1016/j.renene.2020.01.123>
- Wolleswinkel, R. E., de Vries, R., Hoogreef, M. F. M., & Vos, R. (2024). Part I is a new perspective on battery-electric aviation: Reassessment of achievable range. *AIAA Scitech 2024 Forum*. <https://doi.org/10.2514/6.2024-1489>
- ZeroAvia. (2023). Hydrogen-Electric Propulsion for Regional Aircraft: Q400 Feasibility Study. ZeroAvia Technical Report.

Economic Concerns Perpetuating the Occurrence of Type II Diabetes in AI/AN Communities in the United States By Aabha Vadapalli

Abstract

Type II Diabetes (T2D) disproportionately affects American Indian/Alaska Native (AI/AN) communities, with more than 15% of the population impacted, according to the Indian Health Service (Special Diabetes Program for Indians, 2016). Underfunding of the Indian Health Service (IHS) and high levels of underinsurance contribute to this issue, with 14.6% of AI/AN individuals being underinsured (Lee et al., 2021). Funding for AI/AN healthcare continues to decline, as noted by the National Tribal Budget Formulation Workgroup (Torres et al., 2021). Efforts to address these challenges are often misdirected, targeting areas where the need is less critical (Pollak, 2020). As a result, the rising prevalence of T2D in AI/AN communities remains a major health concern, exacerbated by inaccessibility to care, insurance undercoverage, and inadequate healthcare funding.

Keywords

Type II Diabetes, Native Americans, Health Disparity, Insurance, Funding

Introduction

Type II Diabetes (T2D) disproportionately impacts AI/AN communities in the U.S., according to the Indian Health Service (IHS)—the primary AI/AN healthcare provider, it is present in more than 15% of the population (Special Diabetes Program for Indians, 2016). Lack of economic support and financial resources are driving this rate upward. Dr. De Chi Lee MD, a health equity researcher at Da-Yeh University, notes that the IHS is underfunded, receiving thousands of dollars less than similar organizations, and around 14.6% of the AI/AN population is underinsured (Lee et al., 2021). Furthermore, the National Tribal Budget Formulation Workgroup, a tribal financial management group, establishes that funding is only decreasing (Torres et al., 2021). Additionally, as Medical anthropologist from Indiana University Northwest Dr. Margaret Pollak notes, efforts to combat these issues target the wrong regions (Pollak, 2020).

All these factors lead to issues such as the inaccessibility of Medical facilities and basic consistent healthcare. The rising T2D rates in AI/AN communities is a growing problem, primarily perpetuated by insurance undercoverage and underfunding of major organizations.

Underinsurance of AI/AN Population

An economic issue perpetuating T2D rates among Native Americans is the inaccessibility of health insurance. In the U.S., Native Americans are the group that suffers most, even compared to other struggling minorities. According to a report by the national organization U.S. Department of Minority Healthcare, “...14.9 percent of AI/ANs had no health insurance coverage. This compares to non-Hispanic whites by 74.7 percent, 34.3 percent, and 6.3 percent respectively,” highlighting the severe underinsurance of AI/AN, even compared to other

struggling groups (*July 2021 ISSUE BRIEF*, 2024). Of Native Americans, 15% have no access to insurance, in other words, 1 in 7 AI/AN citizens can't get medical care without having to foot an unaffordable bill. Dr. Pranav Patel, a researcher at the University of Toledo College of Pharmacy, further establishes that lacking insurance reduces access to medical resources such as diabetes technology in U.S. minority groups. Meaning, AI/AN—a minority group in the U.S. lacking insurance—can't access key technologies or treatments such as Continuous Glucose Monitoring Devices. (Patel et al., 2024).

High rates of uninsurance are driven by lack of funding in regions with majority of AI/AN population. According to the non-partisan government organization MacPac, which evaluates U.S. health insurance companies, thousands of Native Americans are located outside geographic service areas of IHS, where it is harder to access insurance. Additionally, IHS programs are concentrated in rural areas, while 80% of the native population is in urban areas; it highlights the shockingly disproportionate allocation of funds. (*Advising Congress on Medicaid and CHIP Policy*, 2021). Renowned Physician and Indigenous American equity advocate, Dr. Warne MD MPH expands on this issue highlighting how the set-up of government-funded insurance coverage for Native Americans keeps funds from urban areas, stating that “States are reimbursed with 100% Federal Medical Assistance Percentage for payments made to [IHS service] programs, and therefore, state funds are not used to pay for Medicaid-covered-services... However, the urban Indian health centers are not eligible for 100% Federal Medical Assistance” (Warne & Frizzell, 2014). Essentially, he emphasizes that states aren't reimbursed for funding urban healthcare centers, which means funding doesn't go to those centers. Furthermore, MacPac adds, that under current law, Indian health providers like the IHS can only receive Medicaid payment for services authorized by the state plan—which concentrates on reservations and rural areas because that is what the government supports. This means Medicaid insurance is only accessible to reservations and the rural regions that IHS service areas comprise, but not to centers that the majority of the population can access. Therefore, the federal government's selective reimbursement of state funds, causes state Medicaid program funding to be limited, thus narrowing IHS's bandwidth and Native Americans access to care. (Warne & Frizzell, 2014).

Underfunding AI/AN Health Service Organizations

Another economic issue driving T2D rates upward is the underfunding of AI/AN healthcare funding organizations. Dr. Trula Breuninger— CEO of Native American Connections, an AI/AN health advocacy group—establishes that every branch of the AI/AN healthcare system is underfunded: IHS, Tribal Health, and Urban Indian Health programs. Her analysis suggests that a deficiency in funding of Native American services is driving their lack of healthcare access (Kaplan, 2023).

The underfunding of organizations is notable. According to Dr. Ankit Mangla, a Case Western University Researcher and physician at the Seidman Cancer Hospital, on average IHS spends \$3,332 per patient. In contrast, Medicare spends around \$12,829 per patient, highlighting the gap in funding between the rest of the U.S. and Native American health services (Mangla &

Agarwal, 2023). This situation is only worsening. As the National Tribal Budget Formulation Workgroup establishes—necessary funds are essentially decreasing because the “4-3% annual increase to the [Indian Health Service] budget does not keep pace with year-to-year increases in medical inflation, which is now projected to be 4-10%” (Torres et al., 2021). The consistently increasing raw numbers makes federal funding decrease a frequently overlooked issue; when inflation is taken into account it is evident that IHS funding started at a lower point and is increasing at a slower rate than funding for non-AI/AN citizens—a problem limiting Native Americans healthcare access. Furthermore, the National Indian Council on Aging, a government-approved social service organization explains, that cultural and financial barriers that prevent transportation, the ability to take time off work, etc., make it incredibly tough for Native Americans to find aid for themselves. Thus, given their analysis, it is clear that aid must be brought to Native Americans or presented in a way that makes them aware of it. It must be accessible and administered by a trusted organization (*American Indian Health Disparities*, 2024). However, this is not the case. Not only do federal guidelines limit funding of urban centers, as Dr. Warne establishes, but they also don’t, according to MacPac, support mobile clinics, disease intervention programs, and other community-specific solutions that are brought to regions where natives lack access. These programs would help overcome the cultural and financial barriers that the National Aging Council stated would make accessing aid difficult, as these unfunded programs work on community-specific goals, build trust, and come to the people. Federal funding guidelines cause programs like this to be unsustainable and to peter out. The Tribal Sovereign Leaders report adds that government funding limitations highlight their attention to local issues and minimal cooperation with tribes, stating, “Flexibility for federally-operated health facilities and IHS headquarters to have the authority to adjust programmatic funds across accounts. This will maximize the efficiency and effective use of federal dollars at the local level.” Local control means that resources will be addressed by need, instead of priorities that might not be relevant to immediate health issues—again highlighting how the limitation of community-focused and mobile care makes it difficult for the IHS to tailor its decisions to the needs of the communities it services.

Conclusion

T2D is a growing problem in the AI/AN community that is perpetuated by economic factors, including underinsurance and underfunding of major organizations. There is only one notable effort to combat this issue to date—the Special Diabetes Program for Indians. As explained by UC Irvine Public Health professor and T2D disparity researcher, Luohua Jiang, “SDPI-DP Program was a congressionally mandated initiative designed to prevent diabetes among AI/AN communities...[it’s] one of the largest DPP translational efforts in the U.S., especially among racial/ethnic minority populations.” It aims to fund holistic interventions in each tribe connected to IHS that improve diet, treatment access, and education regarding T2D. The program’s funding has been increasing; however, Jiang notes that while this program has shown early success in reducing overall diabetes prevalence, it is not accessible to disadvantaged

people, and is still excluding those outside reservations and rural areas (Jiang, 2022). Therefore, while this program is a start, future interventions should ensure funds are targeted towards the majority of and most greatly impacted populations.

Works Cited

- Advising Congress on Medicaid and CHIP Policy*. (2021).
<https://www.macpac.gov/wp-content/uploads/2021/02/Medicaids-Role-in-Health-Care-for-American-Indians-and-Alaska-Natives.pdf>
- American Indian Health Disparities*. (2024). National Indian Council on Aging.
<https://www.nicoa.org/elder-resources/health-disparities/>
- July 2021 ISSUE BRIEF 1 Health Insurance Coverage and Access to Care for American Indians and Alaska Natives: Current Trends and Key Challenges KEY POINTS*. (2024).
<https://aspe.hhs.gov/sites/default/files/2021-07/aspe-aian-health-insurance-coverage-ib.pdf>
- Breuninger, T (as cited in Kaplan). (2023, November 1). *Native American health care remains vastly underfunded*. Berkeley Public Health.
<https://publichealth.berkeley.edu/alumni/native-american-health-care-remains-vastly-underfunded>
- Jiang, L., Johnson, A., Pratte, K., Beals, J., Bullock, A., Manson, S. M., & Special Diabetes Program for Indians Diabetes Prevention Program (2018). Long-term Outcomes of Lifestyle Intervention to Prevent Diabetes in American Indian and Alaska Native Communities: The Special Diabetes Program for Indians Diabetes Prevention Program. *Diabetes care*, 41(7), 1462–1470. <https://doi.org/10.2337/dc17-2685>
- Lee, D.-C., Liang, H., & Shi, L. (2021). The convergence of racial and income disparities in health insurance coverage in the United States. *International Journal for Equity in Health*, 20(1). BioMed Central. <https://doi.org/10.1186/s12939-021-01436-z>
- Mangla, A., & Agarwal, N. (2023, May 29). *Clinical Practice Issues In American Indians and Alaska Natives*. PubMed; StatPearls Publishing.
<https://www.ncbi.nlm.nih.gov/books/NBK570601/>
- Maudrie, T. L., Aulandez, K. M. W., O’Keefe, V. M., Whitfield, F. R., Walls, M. L., & Hautala, D. S. (2022). Food Stress and Diabetes-Related Psychosocial Outcomes in American Indian Communities: A Mixed Methods Approach. *Journal of Nutrition Education and Behavior*, 54(12). <https://doi.org/10.1016/j.jneb.2022.06.004>
- Patel, P. M., Thomas, D., Liu, Z., Aldrich-Renner, S., Clemons, M., & Patel, B. V. (2024). Systematic review of disparities in continuous glucose monitoring and insulin pump utilization in the United States: Key themes and evidentiary gaps. *Diabetes Obesity and Metabolism*, 26(10). <https://doi.org/10.1111/dom.15774>
- Pollak, M. (2020). Diabetes Care in an Urban Indigenous American Community: Challenges and Suggestions for the Future. *Midwest Social Sciences Journal*, 23(1), 69–79.
<https://doi.org/10.22543/0796.231.1029>
- Special Diabetes Program for Indians | Fact Sheets*. (2016, October). HHS.gov; Indian Health Service . <https://www.ihs.gov/newsroom/factsheets/diabetes/>

- Torres, A., Joseph, V., & Joseph Jr., A. (2021). *Building Health Equity With Tribal Nations* (pp. 28–54). Tribal Co-Chairs.
<https://www.nihb.org/docs/02072022/FY%202023%20Tribal%20Budget%20Formulation%20Workgroup%20Recommendations%20Vol%201.pdf>
- Warne, D., & Frizzell, L. B. (2014). American Indian Health Policy: Historical Trends and Contemporary Issues. *American Journal of Public Health, 104*(S3), S263–S267.
<https://doi.org/10.2105/ajph.2013.301682>

Adaptive Strategies for Sustaining Corn Yields in Iowa: Assessing the Impact of Climate Change on Precipitation and Temperature Variability By Jiahao Wen

Over the past 10 years, how have variations in precipitation levels and other primary factors influenced corn yields in Iowa, and what are the implications for both local and global food production and societal stability? Given Iowa's critical role as a major corn exporter, how have farmers adapted their agricultural practices to maintain or enhance productivity in response to these environmental changes, and what adaptive strategies could further mitigate these impacts?

Abstract

This paper examines how variations in precipitation and temperature due to climate change have impacted corn yields in Iowa over the past decade. It explores the adaptive strategies employed by Iowa farmers to mitigate these impacts and evaluates their effectiveness. Analysis of practical studies, simulations, and case studies highlights that corn yields have experienced significant variability, with droughts and excessive rainfall leading to annual yield losses of up to 30%. Adaptive practices such as the adoption of drought-resistant corn varieties, precision agriculture technologies, and enhanced irrigation systems have shown promising results, with yield stability improvements of 10-20%. Despite these advancements, the research underscores that without further interventions, Iowa's corn yields could decline by up to 30% by 2050 due to projected temperature increases of 2-4°C and unpredictable rainfall patterns. This paper concludes by recommending additional mitigation strategies, including optimized planting schedules, better water management systems, and international collaboration to ensure the sustainability of corn production in Iowa, making it resilient against the growing threats posed by climate change.

Introduction

Maize is one of the most important food sources for human consumption, contributing over 19% of global calories consumed (FAO, 2009). As the global population continues to rise, crop production must increase by 60% by 2050 to meet growing food demand (Alexandratos & Bruinsma, 2012). Iowa, known as the "Corn State," is a crucial contributor to the United States' agricultural output, accounting for approximately 16.6% of the U.S. maize production (USDA, 2020). However, the state's corn yields are increasingly vulnerable to the effects of climate change, particularly variations in precipitation and temperature. Projections suggest that average temperatures in Iowa could rise by 2-4°C by mid-century, exacerbating the challenges already faced by farmers (USGCRP, 2018). These climate change effects have begun to influence agricultural productivity and increase the risk of damaging crops, posing significant challenges for farmers who must adapt to maintain and increase their maize yields. Without adaptive measures, corn yields in Iowa could decrease by 10-30% by 2050 due to rising temperatures and unpredictable rainfall patterns (USGCRP, 2018; Hatfield et al., 2018). This paper aims to explore

how changes in precipitation levels and other key factors have affected corn yields in Iowa over the past decade. We seek to understand how variations in rainfall and temperature (global anthropogenic climate change) have impacted the amount of corn Iowa farmers can grow and harvest. Investigate how Iowa farmers have adapted their farming practices to cope with these environmental changes and evaluate how effective these strategies have been so far. Finally, I will suggest new ways to help farmers further mitigate the impacts of climate change to ensure that food production remains stable and sustainable. This research will provide valuable insights and recommendations for farmers, policymakers, and researchers.

Climate Change and Corn Yields in Iowa

The key issue affecting corn production in Iowa is the increase in temperature associated with global warming. Xu, Twine, & Girvetz (2016) provide deep analysis and experimental research on how these temperature changes are impacting maize yields. This research reveals that the average temperature in Iowa has increased over recent decades, particularly during the growing season. This warming trend is most noticeable during the summer months typically from mid-June to the end of August—when corn is in the critical stage of development—typically when corn is in the V15 (10-12 days pre-tassel) through R2 (blister) stages. This is because corn plants utilize massive amounts of water each day for “cooling” themselves and for sugar production that allows photosynthesis (Coffman, 1998).

According to (Xu, Twine, & Girvetz, 2016), it was found that for every degree Celsius increase in temperature during this critical period, corn yields can decrease as much as 10% (Fig. 1). This reduction is primarily due to the plant’s diminished ability to photosynthesize effectively under extreme heat waves, leading to reduced kernel -refers to the formation of the edible seed portion of the corn- formation and lower overall yields. This study presents data indicating that the frequency of days with temperatures exceeding 30 °C has increased, a condition that significantly damages the photosynthetic efficiency of maize plants. Furthermore, it highlights that during particularly hot years, such as 2012, corn yields in Iowa were much lower than the long-term average underscoring the vulnerability of maize to rising temperatures, which lead to poor arability.

A University of Nebraska-Lincoln article on heat stress and corn yield impact noted:

Corn is a 'C4 Photosynthesis' plant, making it extremely efficient at capturing light and fixing CO₂ (carbon dioxide) into sugars. However, one drawback of this system is that with such high daytime temperatures, the efficiency of photosynthesis decreases, so the plant makes less sugar to use or store. High nighttime temperatures increase the respiration rate of the plant, causing it to use up or waste sugars for growth and development. This results in the plant making less sugar but using up more than it would during cooler temperatures. “That yield loss potential is notable. If soil moisture is adequate, one day of heat with temperatures in the mid-90s Fahrenheit does not stress the

corn plant. But if this heat continues over four consecutive days, corn yield can drop by 1% for each day above that temperature. The potential yield loss is even greater after the fifth or sixth day. A significant additional detail: Damage caused by extreme heat can be partially mitigated by irrigation or increased precipitation, but not eliminated. Once a kernel is lost because of very hot conditions, the corn plant will not get it back (Anderson, 2023).

The Dual Threat of Excessive Rainfall and Drought

Besides the rising heat temperature, variations in precipitation patterns are another vital factor of concern for Iowa's corn farmers. One study (Rosenzweig, Tubiello, Goldberg, Mills, Bloomfield, 2002) highlights that Iowa state has seen a significant increase in both the frequency and the intensity of heavy rainfall. This is indeed problematic because excessive rainfall could potentially lead to waterlogging when there is too much water in a plant's root zone, which decreases the oxygen available to roots, ultimately stunting the plant's growth (*Rosenzweig, Tubiello, Goldberg, Mills, Bloomfield, 2002*). This study provides evidence that these conditions have become more common in Iowa state. Iowa state has experienced rainfall levels that far exceed historical averages. To demonstrate this significance, the analysis showed that during periods of extreme rainfall, waterlogging increased by 25%, which also leads to corn yield decreases of up to 15% (Rosenzweig, Tubiello, Goldberg, Mills, Bloomfield, 2002).

Moreover, heavy precipitation also aggravates soil erosion and nutrient leaching, both of which will degrade soil quality and reduce the availability of essential nutrients for crops, making the land not arable. This also further contributes to yield declines, particularly in fields that are not protected by conservation practices, such as no-till farming. The article reported that the combination of soil erosion and nutrient loss had led to an estimated 10% decrease in productivity in the most vulnerable areas. (Rosenzweig, Tubiello, Goldberg, Mills, Bloomfield, 2002).

Interestingly, the study (*Li, Guan, Schnitkey, DeLucia, Peng, 2019*) emphasizes that the damage caused by excessive precipitation can be one of the causes of drought. The research compares yield losses from both extreme wet and dry conditions across several U.S. states, including Iowa State. The research finds that both extreme conditions lead to similar magnitudes of yield reduction. The study shows that maize yield can decline up to 40% during years of extreme rainfall or drought (*Li, Guan, Schnitkey, DeLucia, Peng, 2019*). This dual threat from both extreme rainfall and drought underscores the complexity of the challenges that Iowa farmers need to face. While drought conditions can only reduce the availability of water for plant growth, excessive rainfall can cause physical damage to the plants and their growing environment, which significantly makes it difficult for farmers to maintain stable yields.

Data Analysis

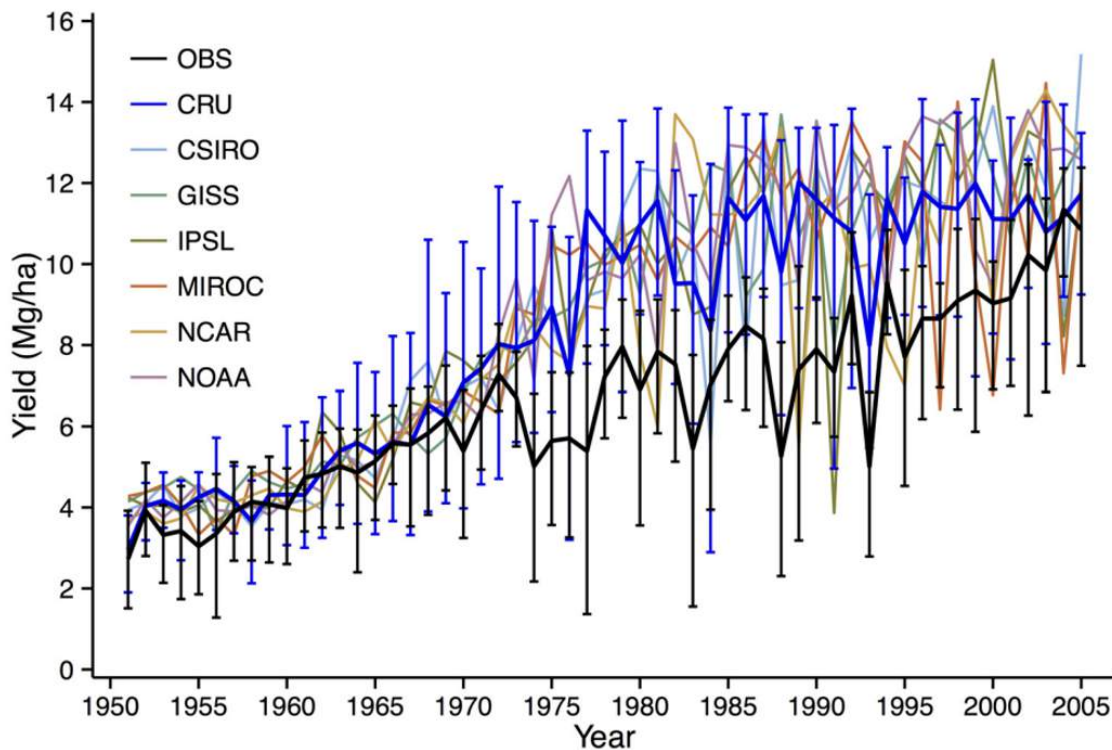


Fig 1. Agr-IBIS simulated maize yield as driven with CRU data and output from all six GCMs from 1951-2005, and maize yield surveyed by USDA (Xu, Twine, & Girvetz, 2016) .

This graph is from (Xu, Twine, & Girvetz, 2016). This graph provides critical insights into the relationship between climate variability and corn production in Iowa (Fig. 1). This graph illustrates simulated maize yields using CRU data and outputs of six different Global Climate Models (GDMs), GCMs, or General Circulation Models, are sophisticated calculation tools used to simulate the Earth's climate system, helping scientists predict future climate conditions by considering factors like atmospheric dynamics, ocean currents, and land surface processes. Which spans from 1951 to 2005. The graph indicates a general upward trend in maize yield over time, this is very likely due to advanced agricultural technology. However, it also shows significant variability in yields, which can be attributed to climate change.

The graph shows that while the average yield has increased, there are noticeable fluctuations if we pay attention to the years associated with extreme events. 1974: There is a significant dip in yield, which may correlate with climate-related stressors. 1988: Another sharp decline in yield, which corresponds to a known drought year in the United States. 1993: There is a noticeable dip, potentially linked to the Midwest flooding that occurred that year. The error bars, representing the range of yield outcomes across different grid cells, further underscore the uncertainty that farmers face due to local climate conditions.

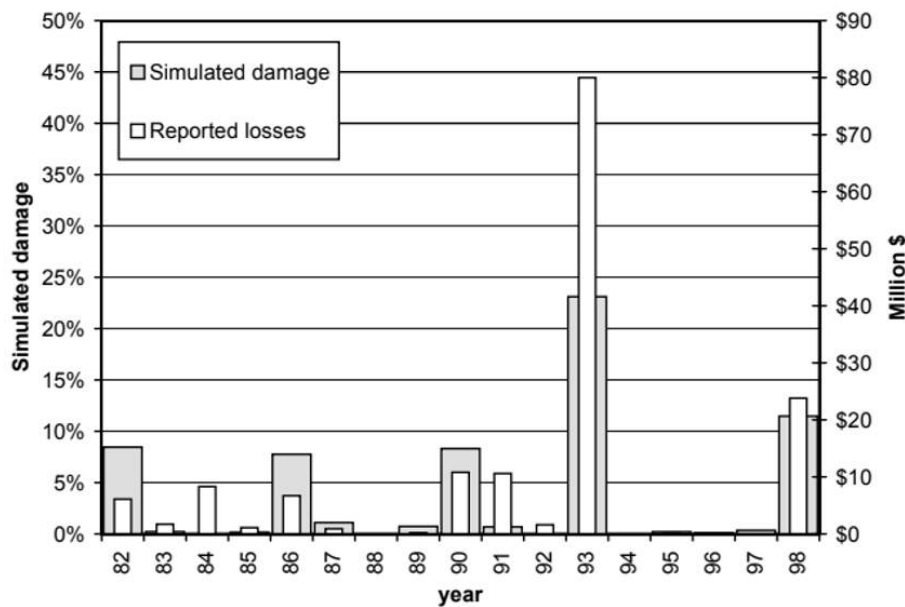


Fig 2. Reported losses in corn production due to excess soil moisture in Polk County, compared to model- simulated crop damage (Rosenzweig, Tubiello, Goldberg, Mills, Bloomfield, 2002).

Figure 2 from the study (Rosenzweig, Tubiello, Goldberg, Mills, Bloomfield, 2002) provides a clear demonstration of how excessive soil moisture due to increasing rainfall had led to significant crop damages and yield loss in Polk country Iowa. The graph compares reported losses in corn production with Model Simulated damage, this reveals a consistent pattern where years with high levels of soil moisture, such as 1993 and 1998 correspond to significant financial losses and damages to corn yields.

The graph shows a direct correlation between excessive rainfall events and the percentage of yield loss, particularly in years where rainfall far exceeded normal levels. For example, in 1993, both simulated and reported losses were at peak, reflecting the severe impact of flooding and waterlogging on corn production. This data underscores the importance of effective water management strategies and adequate drainage infrastructure to mitigate the adverse effects of extreme precipitation events on crop yields.

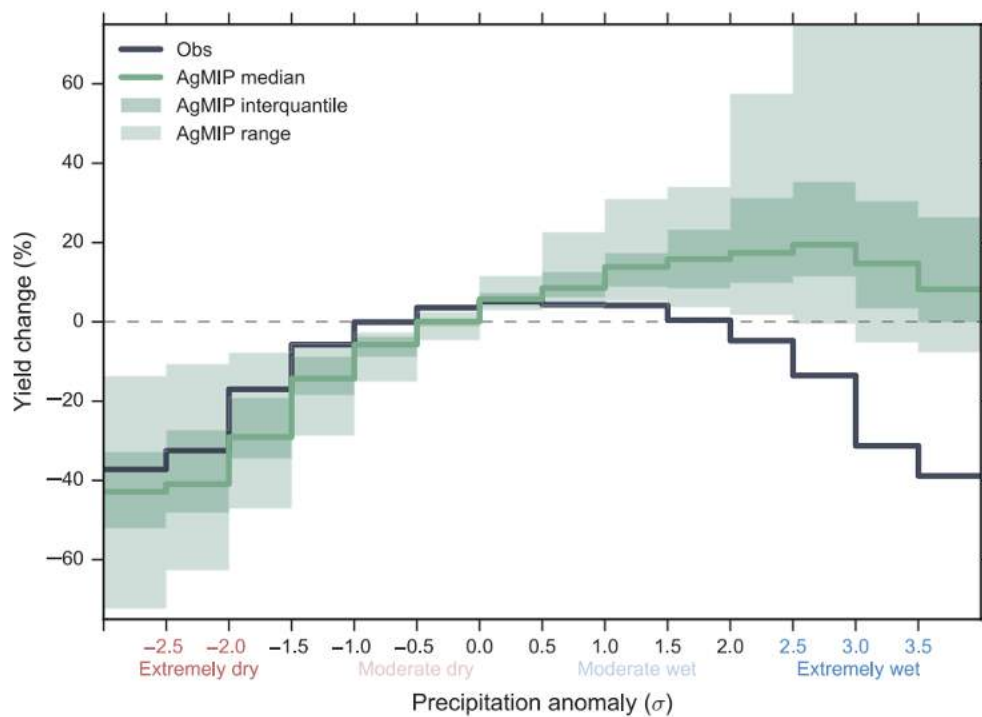


Fig. 3 Maize yields response to precipitation anomaly from simulations of 12 global crop models participated in the AgMIP (Green solid line and shaded area) compared with observed response (black solid line) (Li, Guan, Schnitkey, DeLucia, Peng, 2019).

Figure 3 from (Li, Guan, Schnitkey, DeLucia, and Peng, 2019) offers a detailed analysis of how maize yields respond to precipitation anomalies, this provides insights into the dual threat from both excessive rainfall and drought. The graph compares observations of yield changes predicted by simulations from 12 global crop models. The data showed both extreme dry and extreme wet conditions lead to significant yield losses, with the potential to reduce yields by more than 40% under severe conditions. The graph illustrates that maize yields decrease aggressively when precipitation diverges significantly from normal levels either towards extreme dryness or wetness.

Adaptation Strategies: How Iowa Farmers Have Adapted to Environmental Changes

As climate change continues to alter weather patterns, Iowa's corn farmers are increasingly facing multilateral challenges of maintaining crop yields and changing domestic environmental conditions. These challenges include frequent and intense precipitation, rising temperature and overall climate increase (global warming). To resolve these challenges, Iowa's farmers have adopted key adaptation methods. This section examines these strategies in a detailed analysis. Supported by past accurate data and simulations from recent studies.

Simulated Climate Change Impacts on Corn Yields

This study “Simulated Climate Change Impacts on Corn and Soybean Yields in Buchanan County, Iowa” provides a robust simulation of how future climate scenarios may affect agricultural productivity in the Iowa region, especially focusing on corn yields. With the help of advanced climate models, this study projects that under a high-emission scenario, average temperatures in Iowa could rise by as much as 3°C by 2050. The study’s simulations indicate that for each degree Celsius increase in temperature, corn yields could decrease by 10%, which poses serious challenges for maintaining current levels of production. (Osei, Jafri, Gassman, Gallego, 2023).

Below are the solutions for simulated climate change's impact on corn yields. To effectively address the projected challenges posed by rising temperatures and changing precipitation patterns, a critical approach is essential. The following strategies offer practical solutions that can help Iowa farmers maintain and enhance corn yields in the face of these climate-related pressures.

Adoption of drought-resistant corn varieties

Drought resistant corn varieties are crucial in addressing the yield reduction due to increasing temperatures. These genetically engineered crops are developed to withstand lower water availability by incorporating traits like deeper root systems. By selecting and breeding these corn knells and conducting field tests, farmers can reduce potential yield loss by up to 20% in drought conditions, offering resistance against climate water shortage (Osei, Jafri, Gassman, Gallego, 2023).

Adjusting planting schedules

Adjusting planting schedules accordingly is another key strategy to minimize the effects of extreme heat during those critical growing months, and it plays a crucial role in managing both excessive rainfall and drought conditions. Subsurface drainage systems can prevent waterlogging during heavy rains, while on the other hand-controlled irrigation methods, such as subsurface irrigation help to ensure crops receive adequate water during dry months. Besides, adding soil moisture monitoring tools into these systems allows for more accurate management, which will reduce the risk of over or under-watering. If these measures were being conducted correctly, it would prevent yield losses by up to 30%, especially in extreme weather and rural places.

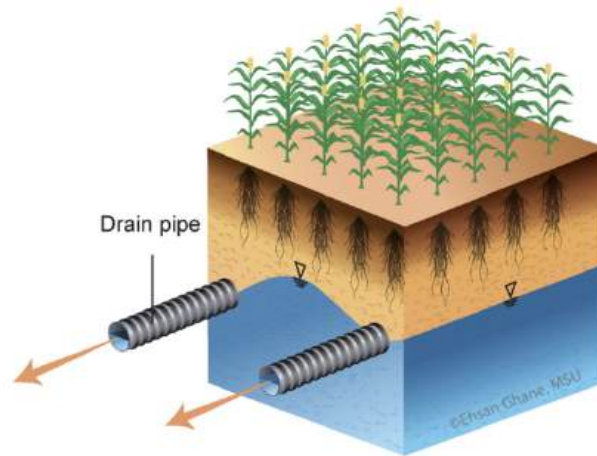


Fig. 4 Diagram of a soil pipe used in subsurface drainage systems, illustrating how excess water is removed from the root zone to improve corn yield and reduce waterlogging (Michigan State University Extension, n.d.).

This is a design illustration presented by Michigan State University on biosystems and agricultural engineering. This illustration was created because surface drainage may be insufficient in poorly drained soils, when the water table is naturally under the ground surface. In these cases, subsurface drainage is installed to remove excess water and to lower the water table. This system should be able to lower the water table from the soil surface to 1ft depth in less than 48 hours after heavy rain (Ghane, 2023).

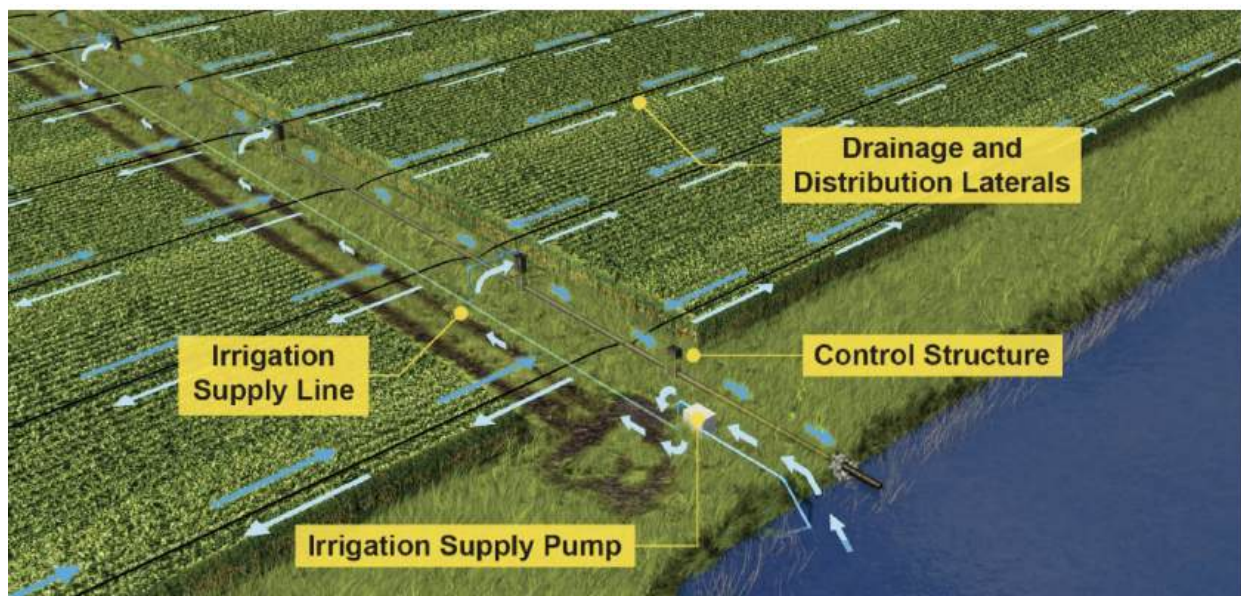


Fig. 5 Diagram of a subsurface irrigation system that maintains optimal soil moisture by distributing water directly to plant root zones, enhancing water efficiency and yield stability (Water Table, n.d.).

This is a figure provided by Prinsco (Water Management Solution). This is similar to a controlled drainage system, however the difference is that in this figure the water supplied to the system is from a ditch, storage pond or lakes that can ensure the water table is always at a sustainable level.

Water can be pumped from a nearby well, ditch or water pond to each control structure, then the water essentially fills up the system to the level of the water control structures. A combination of float pumps and automatic valves will be used to optimize the system. Since water isn't pumped above, sub-irrigation systems can operate more efficiently than sprinkler irrigation. Also, the water is less exposed to the atmosphere on hot and dry summer days, so there could be less evaporation loss compared to other irrigation systems, especially as it is not ideal to feed corn with warm water if the water has been heated by sunlight.

Benefits:

- Higher and less variable yields
- Reduced nutrient loss
- Less energy and water needed to irrigate compared to other irrigation systems

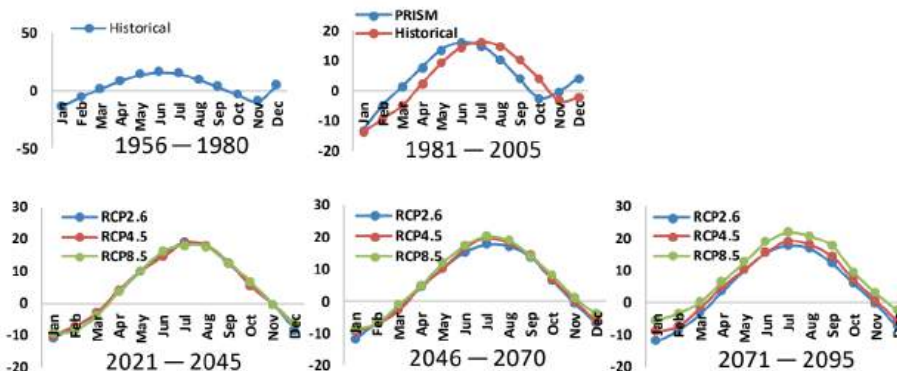


Figure 1. Trends in monthly average minimum temperature (°C) patterns for each climate scenario.

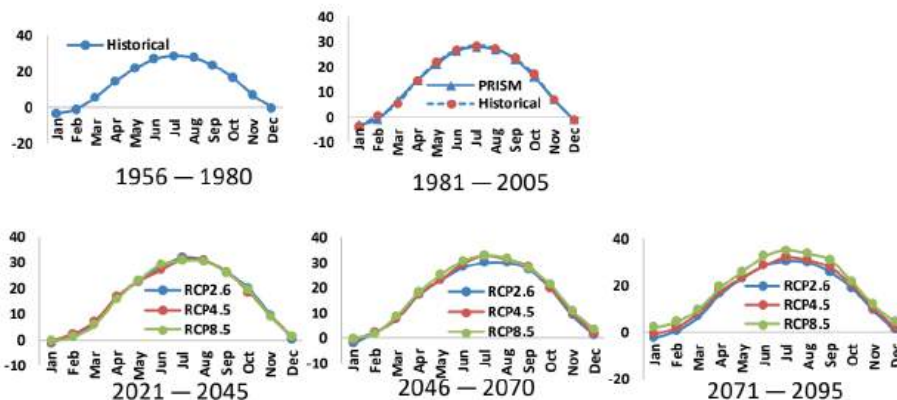


Figure 2. Trends in monthly average maximum temperature (°C) patterns for each climate scenario.

Fig. 6 Simulated climate change impacts on corn and soybean yields in Buchanan County, Iowa, based on model projections assessing yield variability under different climate scenarios (Osei, Jafri, Gassman, & Gallego, n.d.).

The graphs presented in figure 6 offer a detailed projection of temperature trends under various climate scenarios, the data show an expected increase in both minimum and maximum temperature across all future climate scenarios (RCP2.6, RCP4.5, RCP8.5), with the most pronounced changes occurring in the latter half of the 21st century (2046–2095). These trends indicated that the region will experience more frequent and intensive heat during critical growing periods, especially in the summer.

The increase in temperature poses a direct threat, which relates to reduced yields demonstrated in this figure. For example, under the most severe climate scenarios, temperatures are expected to rise, which leads to a 20% decline in corn yields by 2050 if strategies are not implemented. Therefore strategies such as adjusting planting schedules, using drought-resistant crop varieties and implementing advanced water-managed systems to mitigate the effects of climate change and to ensure sustainability (Osei, Jafri, Saleh, Gassman, Gallego, 2023).

Water Management and Drainage Systems

Efficient water management is vital for mitigating the impacts of both excessive rainfall and extreme drought, especially in regions that are prone to heavy precipitation like Iowa state. Subsurface tile drainage mentioned above has been shown to significantly reduce waterlogging, a condition that diminishes oxygen availability to plant roots and can increase the risk of root diseases, ultimately affecting plant growth. For example, the implementation of subsurface tile drainage can decrease the incidence of waterlogging by up to 30%. This will result in a 15 % increase in corn yields during wet years, as it can ensure crops receive sufficient water without overwhelming excess moisture (Helmert, Christianson, Pederson, Brenneman, Iowa State University, 2012).

Drainage Water Management (DWM) systems offer a controlled approach to managing water levels in farm fields, this allows farmers to adjust the water table by varying the drainage intensity based on real-time weather conditions. DWM systems have demonstrated significant improvements in crop resilience, with yield increases reported as high as 15-20% during years of excessive rainfall. The integration of soil moisture sensors, weather stations and real-time data logging equipment allows dynamic adjustments to the water table and ensures that conditions remain optimal throughout the growing season. Additionally, DWM systems enhance water conservation during the dry season by retaining moisture in the soil, reducing the need for supplemental irrigation (Helmert, Christianson, Pederson, Brenneman, Iowa State University, 2012).

Additionally, implementing a *Drainage Water Management* (DWM) system is a complex process that requires detailed planning to maximize its effectiveness. The first step is to conduct a field survey, where the topography is mapped to identify areas that are prone to waterlogging and access soil characteristics such as the type. This stage is very critical because this involves strategic placement and configuration of structures like weirs and gates. This structure is designed to control water tables across the field allowing for more precise management of soil moisture levels. (Helmert, Christianson, Pederson, Brenneman, Iowa State University, 2012).

Real-life application can be seen in Iowa state, where DWM systems have been successfully implemented in farmlands to combat excessive rainfall and fluctuating water patterns. For instance, a farm in southeast Iowa, with support from the Iowa State University Extension and Outreach, installed a DWM system that includes subsurface tile drainage paired with adjustable control panels. This system later helps farmers raise or lower the water table depending on weather conditions and the stages of corn growth (Helmert, Christianson, Pederson, Brenneman, Iowa State University, 2012).

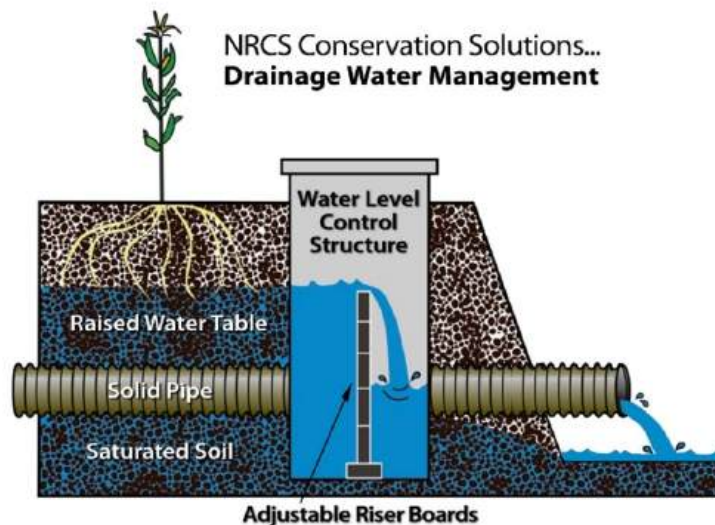


Image courtesy of USDA NRCS.

Fig. 7 Diagram of a drainage water management system illustrating water level control structures for optimized soil moisture regulation (USDA NRCS, n.d.).



NRCS/SWCS Photo by Lynn Betts.

Fig. 8 Photo of a drainage water management system in active use in a field, demonstrating real-world application of water table control structures (Betts, n.d.).

The above figures relate to Drainage Water Management (DWM). “Drainage water management (DWM) helps to control the amount and timing of water leaving agricultural fields through tile lines. DWM works by adding water control structures to existing tile networks to raise or lower the water table to a desired elevation depending on crop needs. DWM works best on flat land, and according to the Iowa Nutrient Reduction Strategy, can reduce nitrate-nitrogen loss through tile lines by an average of 33%.” (Drainage Water Management Facts, Healthy Land Clear Water for life).

Practice benefits

- Manage soil moisture
- Decrease nutrient leaching
- Increase soil carbon (reduced oxidation of soil organic matter)
- Increase yield (estimated 50 to 70 bushels per acre of corn and 20 to 25 bushels per acre of soybeans)

Practice economics

- Estimated to cost \$40 to \$110 per acre (assuming land is flat enough for one structure to control 20 acres)
- Compatible with existing federal and state cost-share programs so farmers who implement DWM can recoup some of their costs

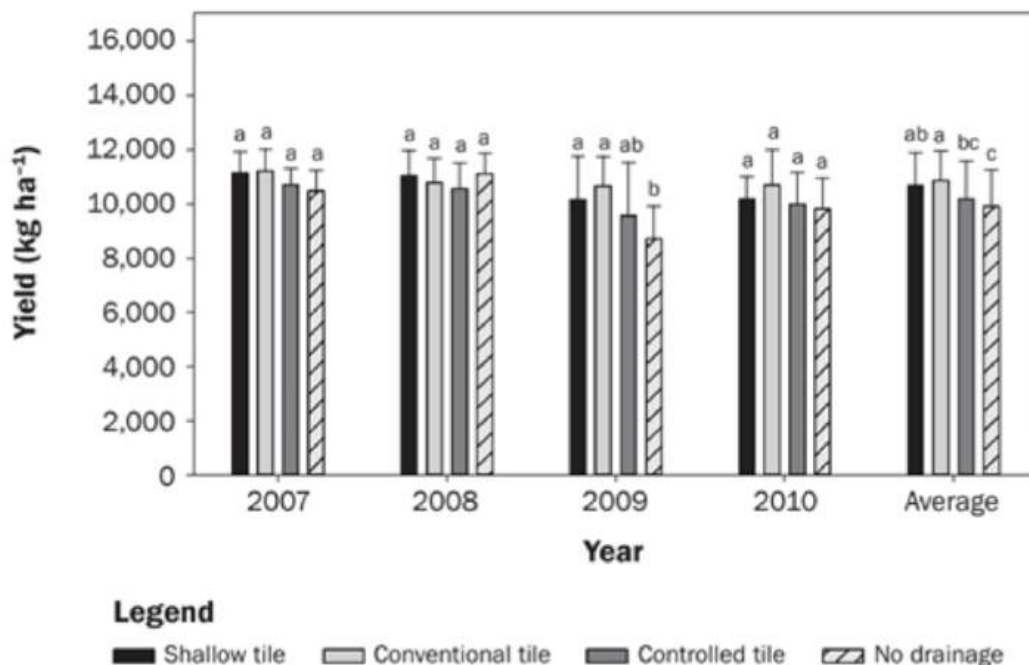


Fig. 9 Yield comparison between shallow tile, conventional tile, controlled tile, and no drainage systems from 2007–2010, illustrating the effects of various drainage practices on corn yields (Helmers, Christianson, Brenneman, Lockett, & Pederson, n.d.).

The graph in Figure 9 shows the importance of efficient water management in enhancing corn yields through different drainage practices. Over the four-year period from 2007 to 2010, fields equipped with subsurface tile drainage systems—both conventional and controlled—consistently produced higher corn yields compared to fields without any drainage system. The data show that the absence of drainage led to noticeably lower yields, especially during wetter growing seasons. By 2009 significant differences in yield became apparent, with fields lacking drainage experiencing noticeably lower yields. The four-year average further underscores the effectiveness of drainage systems, as evidence that both controlled and conventional tile systems led to higher average yields compared to no drainage. This indeed supports the research claim that subsurface tile drainage can lead to a 15% increase in corn yields during wet years by reducing water logging. Yet the real-life application in southeast Iowa mentioned above, where farmers implemented the DWM systems with the help from Iowa State University Extension, demonstrated how these systems effectively manage water levels, optimizing growing conditions and improve crop yields (Helmers, Christianson, Brenneman, Lockett, Pederson, 2012, Fig3).

Future Strategies: New Ways to Mitigate the Impacts of Climate Change

As climate change continues to be a challenge to agriculture, particularly in Iowa's Corn Belt, there is a growing need for innovative strategies to mitigate these impacts. This section explores potential future strategies that can help Iowa farmers sustain corn yields when facing changing environmental conditions. The strategies discussed are informed by recent research on the effects of elevated CO₂ Levels, climate change impacts on global maize production and optimized planting practices.

Leveraging Elevated CO₂ Levels

Rising atmosphere CO₂ levels can have a significant impact on crop yields with the Corn Belt, of which Iowa is part. Elevated CO₂ levels often act as fertilizers, enhancing photosynthesis and water usage. Particularly in C₃ crops like soybeans. However, the benefits for C₄ crops like corn are much more complex. While increasing CO₂ can improve photosynthetic efficiency in corn, these advantages may be decreased by the unfavourable effect of higher temperatures and altered precipitation patterns (Phillips, Lee, Dodson, 1996). To address the climate change problem affecting corn yields, a comprehensive approach is needed. This approach includes breeding CO₂-responsive corn varieties, optimizing irrigation and soil management, adjusting planting schedules, adopting drought-resistant crop varieties, and enhancing soil health through cover cropping. These strategies aim to leverage the benefits of elevated CO₂ while still mitigating the negative impacts of rising temperatures and unpredictable rain patterns (Phillips, Lee, Dodson, 1996).

Breeding CO₂-Responsive Corn Varieties:

Developing corn varieties that are more responsive to the elevated CO₂ concentrations is a critical strategy. These varieties would lead to better photosynthetic efficiency despite extremely stressful conditions such as high temperatures and irregular rainfall. By selecting corn types that allow better carbon fixation and water usage, these varieties could help maintain or increase yields by 5-10% under elevated CO₂ conditions. However, to ensure that the newly modified seeds won't be affected by other climatic stressors, these varieties must be integrated with other adaptive agricultural practices (Phillips, Lee, Dodson, 1996).

Enhancing Soil Health through Cover Cropping

Integrating cover crops into crop rotations improves soil health by reducing erosion, enhancing soil structure and increasing organic matter content. Cover crops help to retain moisture in the soil, making it more resilient to climate variability. Research shows that over time cover cropping can reduce soil erosion by up to 50% and increase soil organic matter by 10% leading to improved water retention and overall crop resilience (Phillips, Lee, Dodson, 1996).

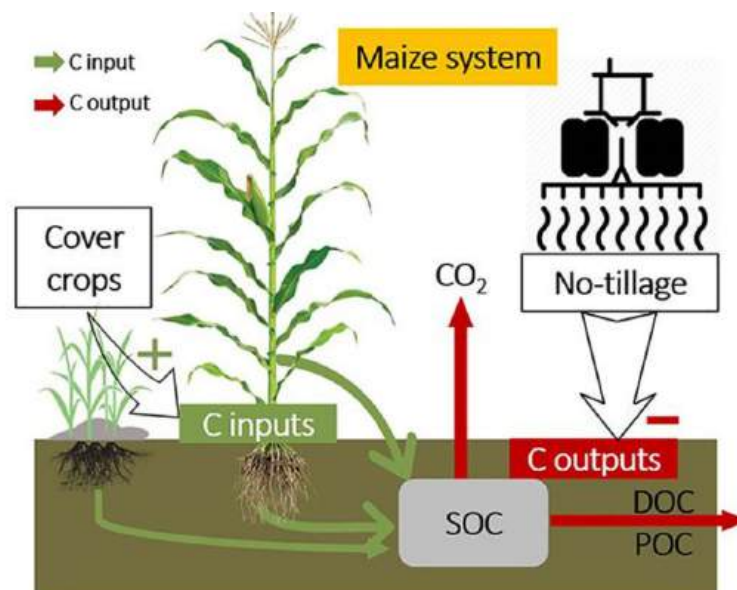


Fig. 10 Diagram of carbon inputs and outputs in a maize system using cover crops and no-tillage practices, illustrating the potential for enhanced soil carbon sequestration and reduced climate impact (University of Kentucky College of Agriculture, Food, and Environment, n.d.).

This figure is provided by the University of Kentucky College of Agriculture, Food and Environment. Cover crops with the combination of no-tillage systems can increase soil organic carbon, which could help to reduce and slow down climate change on farms, according to the study by the University of Kentucky College of Agriculture, Food and Environment. Using an improved agroecosystem such as the one provided in this figure, UK scientists found that

long-term no-tillage together with cover crops would significantly enhance soil carbon sequestration in the context of climate change, and gain more carbon benefits (the University of Kentucky Department of Plant and Soil Sciences- Martin-Gatton College of Agriculture, Food and Environment, Pratt, 2020).

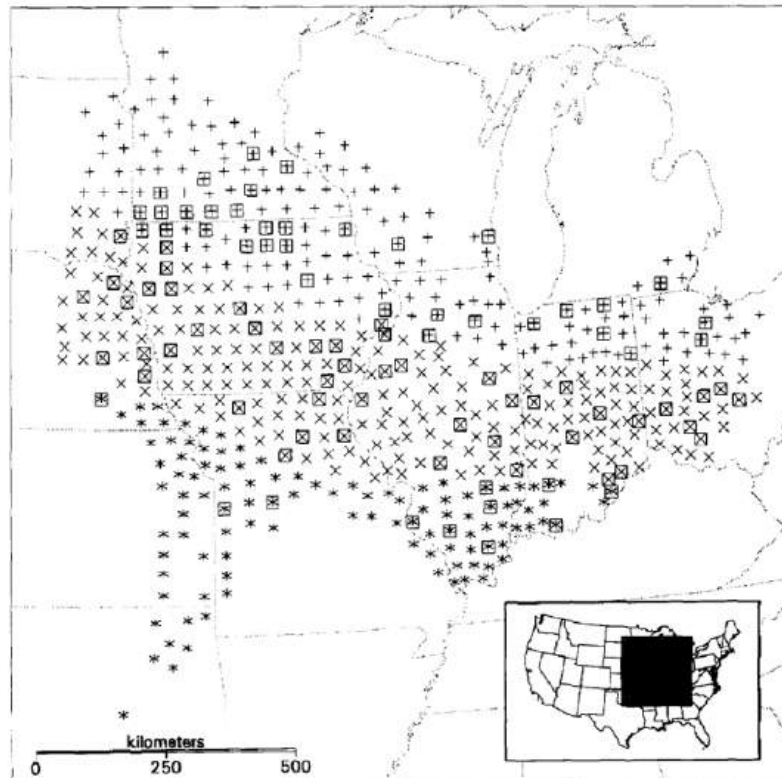


Fig. 11 Map of the spatial distribution of study sites in the U.S. Corn Belt, used to assess regional variability in maize yield sensitivity to climate change and elevated CO₂ conditions (Phillips, Lee, & Dodson, 1996).

Figure 11 illustrates the geographical area in 1987 within the Corn Belt, highlighting regions vulnerable to soil erosion. The map uses various symbols such as crosses, squares, and asterisks to separate different study sites. The concentration of these symbols indicates areas of particular focus of the study.

The purpose of this map is to provide a visual representation to the spatial distribution of these sites across different climatic regions, which is indeed crucial for tailoring crop calendars and to understand the variability in soil erosion and the impact on crop yields within the regions. (Phillips, Lee, Dodson, 1996).

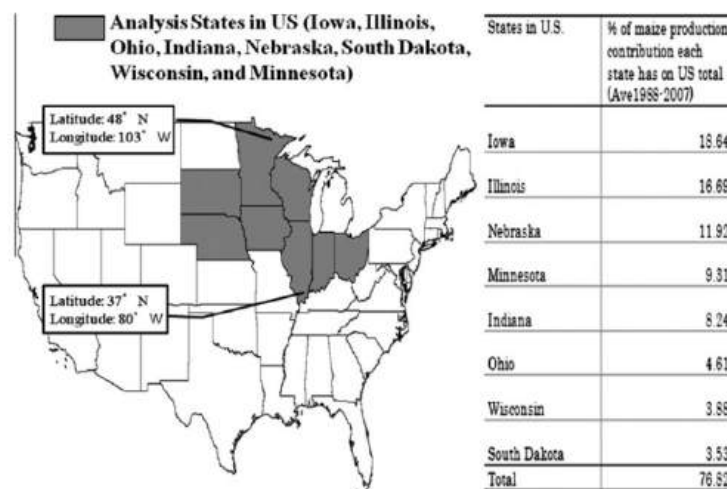
Learning from Global Experiences

Both China and the United States are facing the vital challenges of Corn Yields. In the United States, the projected maize yield could decrease by 10-30% by 2050 if the measure was not taken. This is due to rising temperatures and unpredictable rain patterns. On the other hand,

in China, similar challenges are occurring with yield reduction ranging from 15-25% by mid-century under high emission scenarios.

The impact of extreme weather are particularly influential for both regions. In the U.S, heatwaves during the pollination stage can lead to yield losses up to 50% in Corn Belt regions (Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin). In China, the increasing frequency of drought has also resulted in annual yield loss of 5-10% in the most vulnerable provinces.

Mitigation strategies have shown very tangible results, for example the deployment of drought resistant maize varieties in China has improved yield stability by 10-15%. Moreover, enhanced irrigation to infrastructure has led to a 20% increase in water use efficiency in some regions, this helps to maintain the consistency of maize production despite irregular rainfall. The urgency of these projects highlights the importance of investing in agricultural research and technology implementation. Adaptive practices, such as precision agriculture can optimize inputs and reduce vulnerability to climate stressors by 10-20%(Li, Takahashi, Suzuki, Kaiser, 2010). Solving strategies: In response to the challenges, both China and the United States have implemented strategies to mitigate the negative impacts on maize yields. Here are the strategies implemented by both countries focusing on enhancing crop resilience, improving water management, and optimizing agricultural practices to sustain production in the face of rising temperatures and changing precipitation patterns.



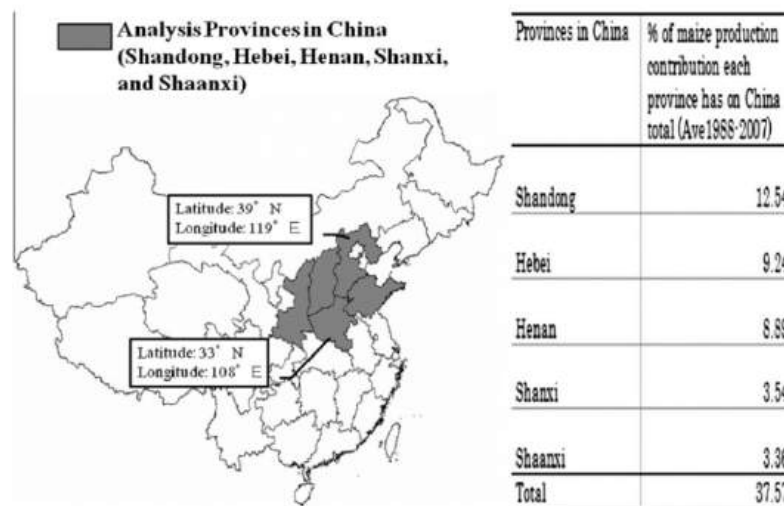


Fig. 12. Maps and tables showing the geographic range of maize production in the Midwestern United States and Middle China, highlighting key agricultural zones vulnerable to climate change impacts (Li, Takahashi, Suzuki, & Kaiser, 2011).

Figure 12 illustrates the geographic range of maize production in both the Midwestern United States and Middle China, focusing on key states and provinces that contribute significantly to national maize yields. In the U.S., Iowa, Illinois, and Nebraska together account for over 40% of the country's maize production, while in China, Shandong and Hebei play a major role.

Both regions face similar challenges due to climate change, with projected maize yield decreases of 10-30% in the U.S. and 15-25% in China by 2050. These figures emphasize the urgency of addressing extreme weather impacts, such as U.S. heatwaves and droughts in China.

To mitigate these challenges, both countries have implemented strategies like developing drought-resistant maize varieties and enhancing irrigation infrastructure. In the U.S., precision irrigation has boosted water-use efficiency by 20-30%, while China's drip irrigation has increased efficiency by up to 40%. Precision agriculture technologies are also helping improve crop resilience in both nations. These efforts underscore the importance of international collaboration in combating climate-related threats to maize production (Li, Takahashi, Suzuki, Kaiser, 2010).

Works Cited

- Abendroth, Lori J., et al. "Yield-Based Corn Planting Date Recommendation Windows for Iowa." Crossmark, 19 May 2017,
crossmark.crossref.org/dialog/?doi=10.2134%2Fcfm2017.02.0015&domain=pdf&date_stamp=2017-05-19.
- Alexandratos, Nikos, and Jelle Bruinsma. "World Agriculture towards 2030/2050." ESA Working Paper No. 12-03 June 2012 Agricultural Development Economics Division Food and Agriculture Organization of the United Nations, 2012,
www.fao.org/fileadmin/templates/esa/Global_persepectives/world_ag_2030_50_2012_rev.pdf.
- Anderson, Bryce. "How Extreme Heat Can Affect Corn Yield." DTN Progressive Farmer, DTN Progressive Farmer, 27 July 2023,
www.dtnpf.com/agriculture/web/ag/blogs/ag-weather-forum/blog-post/2023/07/27/extreme-heat-can-affect-corn-yield.
- Baschea, Andrea D., et al. "Simulating Long-Term Impacts of Cover Crops and Climate Change on Crop Production and Environmental Outcomes in the Midwestern United States." Agriculture, Ecosystems & Environment | Journal | ScienceDirect.Com by Elsevier, 28 Nov. 2015, www.sciencedirect.com/journal/agriculture-ecosystems-and-environment.
- Dai, Panxi, and Benkui Tan. "The Nature of the Arctic Oscillation and Diversity of the Extreme Surface Weather Anomalies It Generates." AMETSOC, American Meteorological Society, 15 July 2017,
journals.ametsoc.org/view/journals/clim/30/14/jcli-d-16-0467.1.xml.
- "Drainage Water Management Facts." Soil and Water Conservation Society, Conservation media library ,
www.swcs.org/resources/conservation-media-library/drainage-water-management-facts. Accessed 22 Mar. 2025.
- Ghane, Ehsan. Agricultural Drainage, Michigan State University, 2023,
www.canr.msu.edu/field_crops/uploads/files/E3370%20AgriculturalDrainage_Rev2.pdf.
- "Gold Tech Note / Water Quality BMP – Denitrifying Bioreactor ..." Prinsco Water Managements Solution , Prinsco,
www.watertable.ag/wp-content/uploads/2021/07/TN-3-013-Water-Quality-BMP-Denitrifying-Bioreactor.pdf. Accessed 22 Mar. 2025.
- Haskett, Jonathan D., et al. "Increase of CO₂ and Climate Change Effects on Iowa Soybean Yield, Simulated Using GL YCIM." Crossmark, Mar. 1997,
crossmark.crossref.org/dialog/?doi=10.2134%2Fagronj1997.00021962008900020004x&domain=pdf&date_stamp=1997-03-01.
- Helmert, Matt, and Greg Brenneman. "Drainage Water Quality Impacts of Various In-Field Nutrient Management Practices: Southeast Research and Demonstration Farm." Iowas State University , June 2017.
- Helmert, Matthew J., et al. "Water Table Response to Drainage Water Management In ..." 495NOV/DEC 2012—VOL. 67, NO. 6JOURNAL OF SOIL AND WATER CONSERVATIONExcessive

- Nitrate-Nitrogen (NO₃-N), as It Relates to Water Quality in the Midwest, Is Tied to Hypoxia in the Gulf of Mexico (Turner and Rabalais 1994). Major Con-Tributing Areas Relative to NO₃-N Loading to the Gulf of Mexico Are the Ohio and Upper Mississippi River Basins, Which from 1997 to 2006 Were Estimated to Contribute 5.9 and 7.2 Kg N Ha⁻¹ y⁻¹ (5.3 and 6.4 Lb N Ac⁻¹ Yr⁻¹) of Riverine Nitrate (NO₃), Respectively (David et al. 2010). Since a Significant Portion of NO₃-N Export Originates from Areas Where Agricultural Subsurface Drainage Systems Are Utilized (David et al. 2010), There Is a Need to Matthew Helmers Is an Associate Professor and Reid Christianson Is an Extension Program Specialist in the Department of Agricultural and Biosystems Engineering, Greg Brenneman Is a eld Specialist in Extension to Agriculture, and Delise Lockett Is a Graduate Student and Carl Pederson Is a Research Associate in the Depart-Ment of Agricultural and Biosystems Engineer-Ing at Iowa State University, Ames, Iowa. Water Table, Drainage, and Yield Response to Drainage Water Management in Southeast Iowa, Nov. 2012, www.researchgate.net/publication/266741051_Water_Table_Response_to_Drainage_Water_Management_in_Southeast_Iowa.
- Houser, Matthew. “Who Framed Climate Change? Identifying the How and Why of Iowa Corn Farmers’ Framing of Climate Change.” Michigan State University, Jan. 2018.
- Kim, Nari, and Lee Yang-Won. “Machine Learning Approaches to Corn Yield Estimation Using Satellite Images and Climate Data: A Case of Iowa State.” Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, 31 Aug. 2016, koreascience.or.kr/article/JAKO201625752504104.page.
- Li, Xiang, et al. “The Impact of Climate Change on Maize Yields in the United States and China.” Agricultural Systems | Journal | ScienceDirect.Com by Elsevier, 16 Dec. 2010, www.sciencedirect.com/journal/agricultural-systems.
- Li, Yan, et al. “Excessive Rainfall Leads to Maize Yield Loss of a Comparable Magnitude to Extreme Drought in the United States.” Crossmark, Wiley, 16 Mar. 2019, crossmark.crossref.org/dialog/?doi=10.1111%2Fgcb.14628&domain=pdf&date_stamp=2019-04-29.
- Osei, Edward, et al. “Simulated Climate Change Impacts on Corn and Soybean Yields in Buchanan County, Iowa.” MDPI, Multidisciplinary Digital Publishing Institute, 22 Jan. 2023, www.mdpi.com/2077-0472/13/2/268.
- Phillips , Donald L., et al. “Sensitivity of the US Corn Belt to Climate Change and Elevated CO₂: I. Corn and Soybean Yields.” Agricultural Systems, Elsevier, 19 Mar. 1999, www.sciencedirect.com/science/article/abs/pii/S0308521X96000145.
- Pratt, Katie. “UK Research: Cover Crops, No-till Could Slow Climate Change.” College News, University of Kentucky Martin - Gatton College of Agriculture, Food and Environment , 14 July 2020, news.ca.uky.edu/article/uk-research-cover-crops-no-till-could-slow-climate-change.
- Rosenzweiga, Cynthia, et al. “Increased Crop Damage in the US from Excess Precipitation under Climate Change.” Pergamon.
- “Soil and Crop Sciences Publications.” Department of Soil and Crop Sciences, 28 Feb. 2025, soilcrop.tamu.edu/publications/.
- Xu, Hong, et al. “Climate Change and Maize Yield in Iowa.” PLOS ONE, Public Library of Science, 2016, journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0156083.

Examining the Effect of Narratives on the Persistent Appeal of the Gold Standard

By Michael Raso

I. Abstract

The gold standard is a striking paradox as it is widely rejected by economists due to its well-documented limitations, but it continues to captivate politicians, commentators, and the general public, revealing a disconnect between economic consensus and popular opinion.. This research examines how four key narratives sustain interest in the gold standard: criticism of fiat money and the Federal Reserve, libertarian/conservative ideology, gold as a store of value, and historical revisionism. Based on Robert Shiller's framework for narrative economics and the idea of consilience, this study analyzes scholarly works, news reports, publications from think tanks, and statements made by influential figures to determine how these narratives perpetuate despite contradicting economic consensus. The findings of this study reveal that these narratives gain traction during economic uncertainty, amplified by celebrities and influential figures who simplify complex monetary concepts for public consumption. The study demonstrates that the gold standard's lasting appeal is due to its capacity to address core issues with inflation, financial security, and government control rather than its economic viability. Although mainly concentrating on U.S. perspectives, this study sheds light on how economic narratives endure and shape public opinion, emphasizing the significance of understanding ideological and cultural forces in discussions of monetary policy. These findings suggest that policymakers should address underlying public concerns about financial stability rather than dismissing gold standard advocacy outright.

II. Introduction

Why do so many advocate for a monetary system abandoned nearly a century ago? Despite unanimous economic rejection, the gold standard still holds public interest (Lewis, 2015).

Although most countries abandoned this system decades ago, the concept of reverting to a currency backed by gold continues to fascinate a diverse group of economists, political figures, and members of the public (Lioudis, n.d.; Cooper et al., 1982). This research dives into the lasting narratives surrounding the gold standard, seeking to understand their persistence despite widespread academic consensus against its reimplementation.

This research applies Shiller's concept of consilience - the integration of insights across multiple fields - to understand why these narratives about the gold standard persist (Shiller, 2019). For example, psychological research on cognitive biases, such as nostalgia or loss aversion, helps explain why people may idealize past monetary systems despite economic evidence against them. Similarly, cultural memory studies reveal how historical narratives and ideological associations shape public perceptions, reinforcing support for the gold standard beyond purely economic reasoning. By examining the intersection of economics, psychology, and cultural memory, we can better understand how narratives about the gold standard continue

to shape public opinion and policy debates. While economic analysis alone might suggest that support for the gold standard should have faded with academic consensus, the persistence of these narratives reveals a more complex interaction between historical memory, ideological beliefs, and economic concerns.

The gold standard, which previously powered the global financial system, was phased away due to its shortcomings, as demonstrated during the Great Depression (Konkel, 2018). Nonetheless, support for its implementation has persisted, often growing in popularity during periods of economic uncertainty. Recent examples include the 2008 financial crisis, the COVID-19 pandemic, and the inflation concerns of the early 2020s (The Royal Mint, n.d.; O'Brien, 2012). This paper contributes to the argument that continual advocacy for the gold standard goes beyond just economic considerations. It reflects deeper ideological beliefs and reactions to perceived shortcomings in modern monetary systems.

To better understand the enduring appeal of the gold standard, it is essential to examine the four key narratives that sustain interest in its revival. First, critics of fiat money and the Federal Reserve claim that modern monetary policies cause inflation and economic instability, making gold a more reliable alternative (Antoni, 2023; Scott, 2023). This skepticism is frequently associated with libertarian and conservative ideology, with supporters viewing gold as a tool for limiting government intervention and restoring fiscal discipline (Winfrey, 2024). Furthermore, gold's reputation as a store of value and a hedge against uncertainty contributes to its appeal, especially during times of financial turmoil (Duggan, 2024; FasterCapital, 2024). Finally, historical revisionism comes into play, as supporters romanticize past gold-backed economies and argue that returning to such a system would bring greater stability and transparency (Forbes, 2024; Maley, 2023).

Despite the widespread academic rejection of the gold standard as a feasible economic policy, these four narratives continue to resonate, especially in times of financial and economic instability (Shiller, 2020; Buckfire, 2024). Advocates often turn to gold not just as a critique of modern economic policies, but as a symbol of a return to perceived fiscal discipline and transparency (Domitrovic, 2019; McMaken, 2024). This fascination endures because the gold standard represents more than a monetary framework. Instead, it is an ideological stance that combines nostalgia for a simpler economic era with skepticism toward government control over money (Paul, 2024; Schiff, n.d.). If concerns about inflation, currency devaluation, and central bank policies persist, the appeal of the gold standard will appear as well.

While the existing literature has extensively examined the economic impracticality of the gold standard, there is still a gap in understanding the ideological and cultural drivers that support its advocacy. This paper addresses that gap by investigating the underlying motivations that drive the gold standard's continued appeal. This helps to provide a more nuanced understanding of how ideological beliefs and historical narratives shape contemporary monetary policy debates. Understanding these motivations is critical for addressing the larger economic and ideological debates that continue to shape monetary policy discussions today (Mueller, 2015).

III.Literature Review

Overview of Key Themes

This research review revolves around four core themes that investigate the relationship between economic narratives, psychological biases, and their effects on financial systems. First, we look at fundamental principles in story economics, emphasizing the importance of stories in affecting economic behavior and decision-making. Next, we look at the psychological biases that make people vulnerable to specific narratives, such as confirmation bias. The third theme focuses on the phenomenon of narrative contagion, analyzing how stories spread and gain traction within societies. Finally, we explore specific narratives surrounding the gold standard, tracing their historical evolution and enduring influence on economic thought. Together, these themes provide a comprehensive framework for understanding the power of narratives in economics.

Foundational Concepts

Narratives are more than just stories - they're powerful tools that shape how we see the world and act within it. According to the Oxford English Dictionary, a narrative is "a story or representation used to give an explanatory or justificatory account of a society, period, etc." (Oxford, 2024). But it's not just any story. A narrative is a particular type of story that promotes a belief, point of view, or a set of values (Shiller, 2019). Therefore, a narrative also emphasizes its "significance to the receiver" (Shiller, 2019 p. 36), and by doing so, they influence consumer behavior and investment trends. Narrative economics, then, is the study of the impact of these narratives on economic matters.

While some may think that it is possible to 'detect' false narratives, it is extremely difficult to do so. Human emotion can act as a barrier that hinders our ability to uncover the truth within narratives. As Shiller puts it "narratives thrive on attachment: human interest, identity, and patriotism" (Shiller, 2019 p. 100). Humans are naturally biased towards information that they agree with, even if it may not be factually correct.

The Role of Celebrities and Psychological Biases

In addition to emotion, psychological biases influence how narratives, particularly economic ones, are perceived and shared, especially when celebrities are involved (Shiller, 2019, p. 51). Celebrities and popular figures frequently act as narrative amplifiers, using their platforms to shape public opinion and societal beliefs. For example, during the late 2010s cryptocurrency boom, figures such as Elon Musk and celebrity endorsements from Kim Kardashian were critical in driving public interest and investment in cryptocurrencies such as Bitcoin and Ethereum. Musk's tweets about Dogecoin caused significant price increases, demonstrating the bandwagon effect, in which people adopted investment behaviors solely because they were endorsed by these well-known figures, often without fully understanding the risks involved. This is an example of narrative contagion, which is the rapid spread of ideas, sometimes regardless of their truthfulness. Media cycles, which have the power to magnify storylines and instill a sense of

urgency or inevitability surrounding particular investments - as is the case with both gold and cryptocurrencies - further fuel this epidemic. While celebrity endorsements are a powerful driver, media coverage also plays a crucial role in shaping public discourse and influencing investment decisions.

Understanding Narrative Through Consilience

Understanding these complex interactions in economics, psychology, and media emphasizes the significance of consilience - the integration of knowledge from various disciplines to achieve a more comprehensive understanding of a phenomena (Shiller, 2019, p. 12). Consilience is seen in the development of public health strategies, which combine epidemiology, behavioral psychology, and data science to track disease spread, understand vaccination hesitancy, and create effective communication campaigns to enhance herd immunity.

According to Shiller, understanding how narratives shape economic beliefs requires the integration of knowledge from multiple fields. By combining insights from various fields, we can better understand how narratives gain traction, influence consumer behavior, and impact investment trends. This shared understanding paints a more complete picture of the forces that shape our economic reality.

Narrative economics expands on behavioral economics by investigating how cognitive biases interact with contagious stories, while simultaneously challenging classical economics by stressing the irrational, emotional drivers of economic decisions. By combining findings from psychology, sociology, and media studies, it provides a more sophisticated understanding of how narratives influence beliefs, actions, and market trends. This interdisciplinary approach emphasizes the need of consilience in identifying the various mechanisms that drive economic phenomena.

Understanding the continued appeal of the gold standard highlights the role of consilience, which Shiller describes as the unification of insights from various disciplines (Shiller, 2019 p. 14). By examining the gold standard through both historical and economic lenses, we can better understand why this concept, despite being outdated, remains influential today. The gold standard is supported by narratives that evoke a sense of stability rooted in history, appealing to economic principles and collective memory. This interdisciplinary perspective helps explain how historical significance and economic arguments combine to support the gold standard's relevance in public and policy debates, showing how past and present are connected through these compelling narratives.

Narrative Constellations and Contagion

In studying why the gold standard remains popular, it's important to look at the ideas of narrative constellations and contagion. Narrative constellations are networks of related stories that support a belief system, while narrative contagion is how those stories spread through populations. Think of constellations as the content (like interconnected stories about gold's stability and government control) and contagion as the transmission method (how these ideas

spread through media and influential voices) (Shiller, 2019 p. 28). For example, libertarian ideology aligns with stories that focus on fiscal responsibility and the stability that a gold-backed currency can provide. Meanwhile, conservative groups might connect with narratives that view the gold standard to protect against government control and monetary manipulation. These connections to specific stories help strengthen group identity and create a sense of belonging, allowing these ideas to persist even though most economists reject the gold standard today.

Narrative contagion explains how these different stories spread within and between groups, especially during tough economic times (Shiller, 2019 p. 18). When worries about inflation and financial stability arise, people are more likely to find the gold standard appealing. These stories travel quickly through social media, political discussions, and endorsements from influential figures, making them more accepted by the public. For instance, when celebrities or respected leaders support the gold standard, they can sway opinions and make the idea more relatable. This interaction between narrative constellations and contagion shows how various stories resonate with different groups, highlighting the emotional and ideological reasons behind the ongoing interest in the gold standard in today's economic debates.

The Gold Standard and Narrative Economics

The discussion of the gold standard today is particularly rooted in narrative economics. The gold standard narrative continues despite being outdated, as it evokes themes of financial security and trust in tangible assets. Such stories circulate through word of mouth and media, showing how emotionally resonant narratives can sustain outdated economic concepts. This paper studies four key narratives that relate to the continued prevalence of the gold standard:

1. Critique of Fiat Money and the Federal Reserve: This narrative spreads through simplified stories about government overreach and monetary manipulation, gaining traction particularly during times of inflation or economic uncertainty. For example, during the stagflation crisis of the 1970s, soaring inflation and sluggish growth fueled distrust in fiat currency, leading to renewed calls for a return to the gold standard. Its effectiveness depends on making complex monetary policy understandable through relatable themes like trust and control (Antoni, 2023; Scott, 2023).
2. Libertarian/Conservative Ideology: The gold standard provides a compelling narrative framework for broader ideological beliefs about government power and human liberty. This story constellation combines with existing myths about free markets and limited government, which increases its attractiveness to specific political groups (Winfree, 2024).
3. Gold as a Store of Value and a Hedge Against Uncertainty: This narrative demonstrates how historical associations can generate compelling economic stories that withstand contradicting data. The emotional resonance of gold's supposed permanence and stability makes this narrative especially widespread during financial crises (Duggan, 2024; FasterCapital, 2024).

4. Historical Revisionism: This narrative shows how selective interpretation of past events can shape present economic beliefs. By portraying the gold standard era as a "golden age," this narrative taps into nostalgia and simplifies complex historical realities into persuasive stories about past prosperity (Forbes, 2024; Maley, 2023).

When celebrities such as Robert Kiyosaki (author of *Rich Dad, Poor Dad*) endorse the gold standard, they lend credibility to these narratives, as shown in this paper, potentially swaying public opinion and attracting followers that were disengaged with the topic (Kiyosaki, 2024). Their influence can make the idea of a gold-backed currency seem more appealing, especially during times of economic uncertainty. This is because economic uncertainty heightens the appeal of a gold-backed currency as it symbolizes security and stability. During financial crises, like the 2008 downturn, people flock to tangible assets like gold, seen as a reliable store of value (Gold Price, 2008). The gold standard offers the promise of stability, making it emotionally appealing in volatile times. Therefore, their endorsement simplifies complex economic issues, allowing audiences to connect emotionally with the idea of returning to a gold standard.

IV.Methods

The study draws on a wide range of sources, including academic articles, books, think tank publications, media reports, and public statements from influential figures. Key sources were selected based on their relevance to the topic, credibility, and the ability to offer diverse perspectives on the gold standard and narrative economics. They included:

1. Scholarly works (especially Robert Shiller's): Shiller's framework, in particular, influenced the identification of rhetorical strategies and assumptions in narratives about the gold standard. His focus on how narratives shape economic beliefs guided the analysis of how these narratives influence public opinion and investment behavior. These provide the theoretical framework of narrative economics and objective analysis of the gold standard's history and economic effects. This foundation allows the study to assess the narratives against established economic principles and historical evidence.
2. Financial news and social media (e.g., Forbes, Schiff, Paul): This data shows how influential figures spread and simplify complex economic ideas for the public, demonstrating how narratives are disseminated and amplified. These sources highlight the role of public figures in shaping perceptions of the gold standard.
3. Think tank publications (e.g., Heritage, Cato, Mises): Think tanks often represent specific ideological viewpoints. These publications are included to analyze how libertarian and conservative ideologies contribute to pro-gold standard narratives, complementing academic sources by revealing the role of organized advocacy in shaping these discussions. By examining these sources, the study can analyze how these narratives are constructed and promoted within ideological frameworks.

Synthesis of Information and Argumentation

The study synthesizes information from these diverse sources to identify and analyze the four key narratives that sustain interest in the gold standard:

1. **Critique of Fiat Money and the Federal Reserve:** This narrative is examined through the lens of public criticism of modern monetary systems, specifically the role of central banks such as the Federal Reserve in managing inflation and economic stability.
2. **Libertarian/Conservative Ideology:** The study looks at how libertarian and conservative ideologies support the gold standard, emphasizing limited government intervention and fiscal discipline.
3. **Gold as a Store of Value and a Hedge Against Uncertainty:** This narrative explores the historical and continuing perception of gold as a haven during times of economic instability.
4. **Historical Revisionism:** The study investigates how nostalgic reinterpretations of the gold standard era ignore its flaws while portraying it as a "golden age" of economic stability.

By putting these sources in conversation with one another, the study identifies common themes and arguments that contribute to the enduring appeal of the gold standard.

Development of a Theory

The study builds on Robert Shiller's theory of narrative economics, which holds those narratives—a specific type of story that promotes a belief, point of view, or set of values—play an important role in shaping economic behavior and policy discussions. The study expands on this theory by applying it to the specific case of the gold standard, demonstrating how narratives can sustain interest in an outdated economic policy despite overwhelming academic opposition.

Specific Design and Procedures

The research design, based on the analysis of 38 sources, involves a qualitative content analysis of the identified narratives.

The first step was data collection. Data collection involved gathering relevant materials from academic journals, media outlets, think tank publications, and public statements made by influential figures.

The next step was categorization. The collected data were classified and sorted into the four key narratives identified in the literature review: Critique of Fiat Money and the Federal Reserve, Libertarian/Conservative Ideology, Gold as a Store of Value and a Hedge Against Uncertainty, and Historical Revisionism. This deductive coding approach, based on predefined

categories derived from the literature, ensured consistency and allowed for a focused analysis of how each narrative contributes to the gold standard's appeal.

The third step was narrative analysis. Each narrative was thoroughly examined through manual coding and close reading using a predefined framework based on the literature, while also allowing for the emergence of new themes through inductive coding. This approach aimed to balance a structured analysis with the flexibility to capture unanticipated nuances in the data. The analysis focused on identifying key arguments, themes, rhetorical strategies, and underlying assumptions employed by gold standard proponents. Detailed notes and summaries documented the findings, ensuring transparency and allowing for a rigorous examination of how each narrative frames the debate.

The last step was comparative analysis. The four narratives were compared based on predefined themes to identify commonalities, contradictions, and how they appeal to people's bias. The analysis concentrated on topics such as attitudes toward government control, fear of loss and economic uncertainty, nostalgia for the past, and ideological influences. These themes assisted in determining how each narrative addresses concerns about government intervention in the economy, financial instability, and idealized views of the past, particularly the gold standard era. For instance, overlaps were found between the Libertarian/Conservative Ideology and the Critique of Fiat Money narratives, as both express skepticism about government control over money and promote gold as a more stable alternative. This analysis offers insight on how these narratives reinforce and challenge one another, demonstrating how they contribute to the gold standard's long-standing appeal, particularly during times of economic uncertainty.

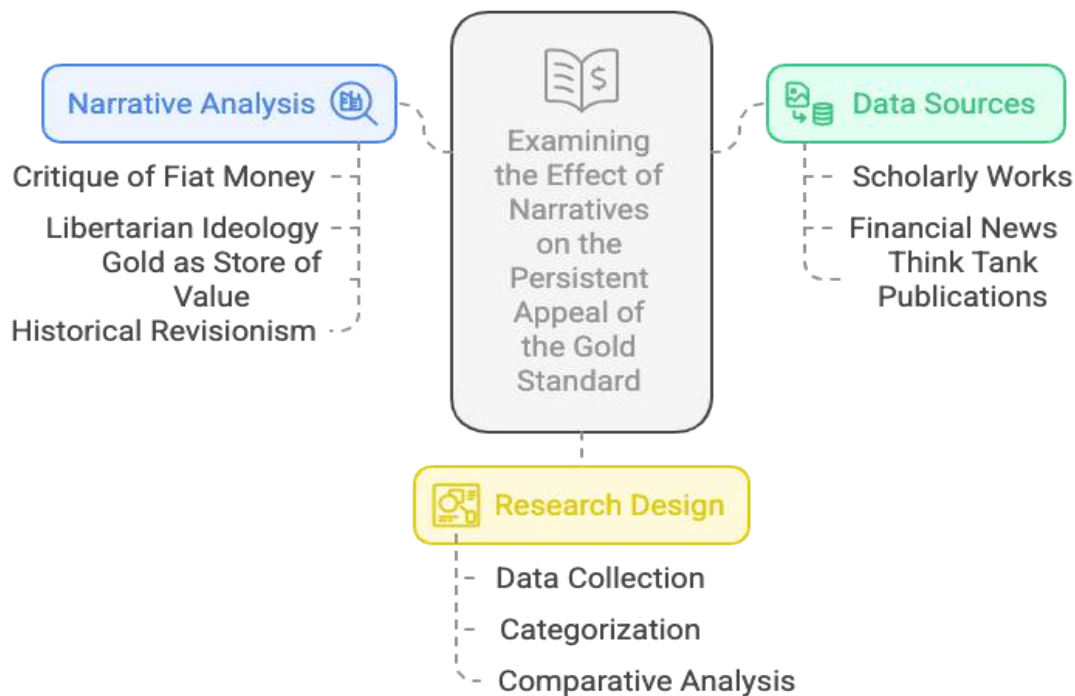


Figure 1: Research methodology diagram showing the three main components of the study: narrative analysis of gold standard arguments, data sources used, and the research design process.

Conclusion

The techniques used in this study, particularly the focus on narrative analysis and the triangulation of diverse data sources, provided a framework for understanding the persistent appeal of the gold standard beyond its economic merits. By examining the four key narratives: critique of fiat money, libertarian/conservative ideology, gold as a store of value, and historical revisionism, the study shed light on the ideological, cultural, and psychological factors that contribute to the gold standard's continued resonance in public discourse and policy debates. This multi-faceted approach, grounded in narrative economics, allowed for a deeper understanding of how these narratives are constructed, disseminated, and received by different audiences. The study contributes to the academic literature by demonstrating the power of narratives to shape economic beliefs and behaviors, even in the face of contrary evidence. Moreover, it offers broader insights into the role of narratives in economic policy and public opinion, highlighting the importance of understanding the non-economic drivers of economic decision-making. While the qualitative nature of the study introduces potential subjectivity, this was mitigated through the development of a pre-defined coding framework, data triangulation, and reflexive consideration of researcher biases. Further research employing quantitative methods could strengthen these findings and explore the impact of these narratives on specific economic outcomes.

V.Results

This section analyzes four key narratives that maintain the gold standard's appeal, examining their transmission mechanisms and impact on public opinion.

Narrative 1: Critique of Fiat Money and the Federal Reserve

Figure 2 illustrates key events and public reactions that fueled criticism of the Federal Reserve, highlighting the pivotal role of monetary policy decisions in shaping this narrative.

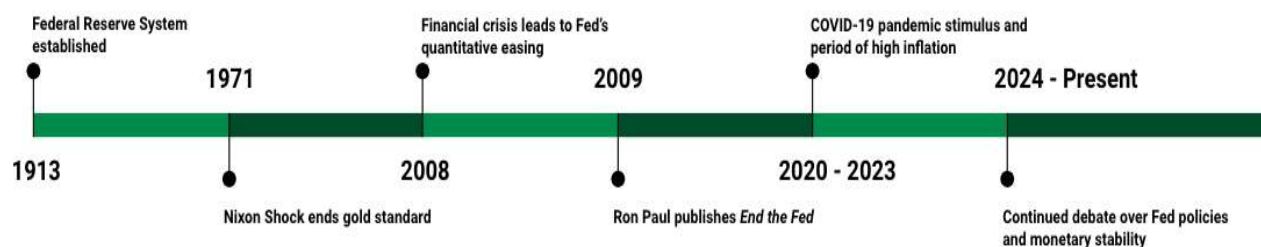


Figure 2: Timeline of Key Events Fueling Federal Reserve Criticism

The Critique of Fiat Money and the Federal Reserve narrative is defined through the arguments that fiat money, which is currency declared legal tender by a government but not backed by a physical commodity, is inherently unstable and susceptible to inflation, devaluation,

and economic mismanagement (O'Brien, 2012). This criticism is also aimed at central banks, particularly the Federal Reserve in the U.S., which is accused of perpetuating these issues through its monetary policies, such as excessive money printing, low interest rates, and currency manipulation (Antoni, 2023; Scott, 2023). Advocates of this narrative propose that returning to the gold standard, where currency value is directly tied to the value of gold, would establish a more stable and reliable economic system, prevent inflation, and safeguard individual savings and financial independence from government interference (Schoenholtz, 2017). This narrative gains simplification through prominent media figures and financial celebrities including like Jim Grant (founder of Grant's Interest Rate Observer), Steve Forbes (Editor in Chief of Forbes), Peter Schiff (financial commentator), and Kevin Vanderboxx (financial analyst), who frequently appear in mainstream media to criticize the Federal Reserve and advocate for a return to the gold standard (Vandenboss, 2024; Forbes, 2024; Maley, 2023).

While the critique of fiat currency and the Federal Reserve often focuses on the perceived technical failures of the system, it also resonates strongly with specific political ideologies. This connection becomes evident in Narrative 2: Libertarian/Conservative Ideology, which frames the gold standard as a crucial tool for limiting government power and promoting individual liberty.

Narrative 2: Libertarian/Conservative Ideology

Figure 3 illustrates a timeline of key events and the evolution of Libertarian and Conservative Monetary Policy Advocacy, highlighting the influence of organizations like the Cato and Mises Institutes, the role of figures like Ron Paul, and the impact of economic events in shaping this narrative. It supports the findings by visually demonstrating the continued and organized advocacy for a gold standard within these ideological frameworks.

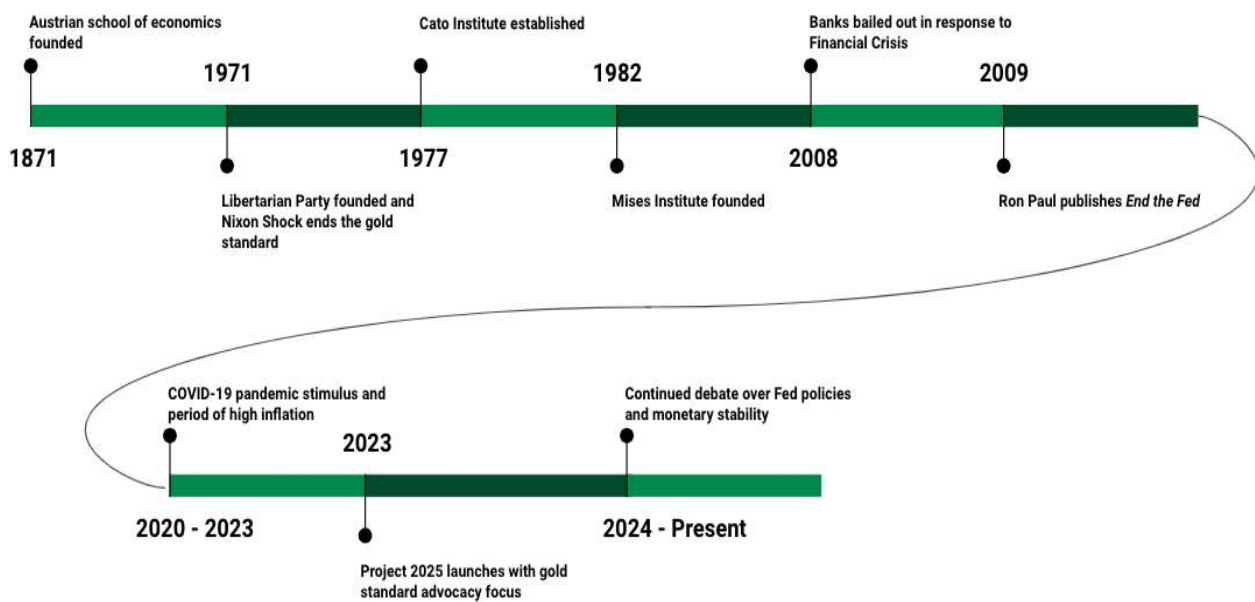


Figure 3: Evolution of Libertarian and Conservative Monetary Policy Advocacy

The Libertarian/Conservative Ideology narrative claims that a return to the gold standard is essential for limiting government intervention, promoting free-market principles, and preserving individual economic freedom (Paul, 2024). This criticism is aimed at modern monetary policies, particularly those of central banks like the Federal Reserve, which are accused of promoting inflation, economic instability, and government overreach (Murphy, 2024). Advocates propose that a gold-backed currency would naturally restrict government power by limiting its ability to print money and manipulate currency, thus allowing economic systems to be governed by market forces rather than centralized control (Rothbard, 2024). This ideology gains substantial influence through notable libertarian and conservative voices including Ron Paul (former U.S. congressman and libertarian leader) and Steve Forbes, along with conservative initiatives such as Project 2025, which is actively advocating for a return to the gold standard (Winfrey, 2024 ; McMaken, 2024). This narrative is further amplified through large organizations such as the Heritage Foundation, the Cato Institute, the American Institute for Economic Research (AIER) and the Mises Institute with campaigns such as Project 2025 (Heritage Foundation), Mises Institute's educational efforts through regularly organized events to discuss monetary policy, and Cato Institute's and American's Institute for Economic Research's publications and frequent media discussions promoting the gold standard (Vandenboss, 2024; Forbes, 2024; Mueller, 2015). These influential figures and organizations shape public opinion through media appearances, social media posts, books, and think tank publications that translate complex monetary theories into simple, compelling arguments for the average American. The libertarian/conservative arguments for the gold standard, based on skepticism of government intervention, frequently meet with a more pragmatic concern: safeguarding wealth in uncertain times. This leads directly into Narrative 3: Gold as a Store of Value and a Hedge Against Uncertainty, which focuses on gold's perceived intrinsic value and past performance during economic downturns.

Narrative 3 Gold as a Store of Value and a Hedge Against Uncertainty

Figure 4 illustrates gold's price peaks during periods of economic instability (inflation, financial crises, pandemics, etc.), supporting the narrative of gold as a safe haven.

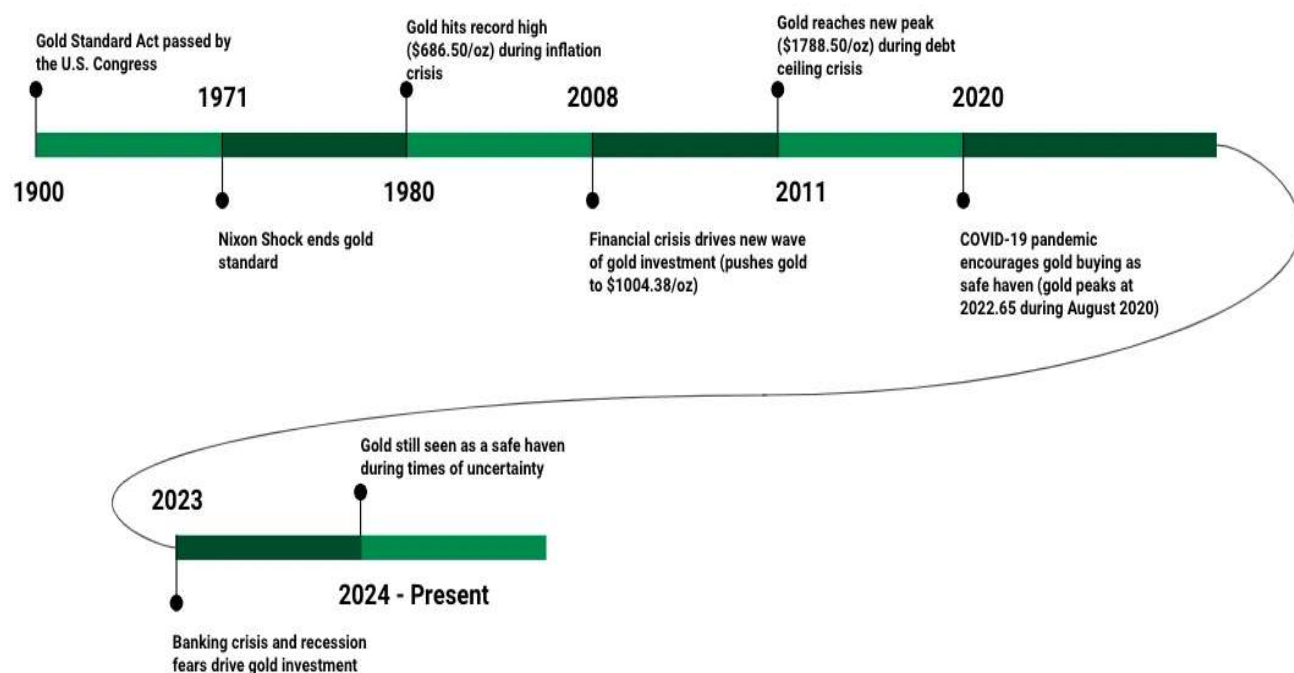


Figure 4: Gold's Performance During Economic Crises

The argument presented in the Gold as a Store of Value narrative is that gold serves as a reliable hedge against economic uncertainty, inflation, and currency devaluation (Gold Price, 2002). This perspective emphasizes gold's historical and ongoing role as a stable asset, contrasting it with fiat money, which can be printed at will by governments and central banks (Duggan, 2024). Advocates propose that gold's intrinsic value and finite nature make it a superior store of wealth, particularly during times of economic turmoil or when trust in fiat currencies wanes (FasterCapital, 2024 ; Bromberg, 2024). Psychological biases, such as loss aversion and the endowment effect, can further amplify this perceived security. This narrative gains traction through well-known financial advisors and market commentators like Robert Kiyosaki (author of *Rich Dad Poor Dad*), Jim Rogers (co-founder of the Quantum Fund), and Peter Schiff (CEO of Euro Pacific Capital), who frequently appear on financial news networks and use social media platforms to prompt gold investments as protection against economic instability. These appearances by these famous influencers with large followings convince everyday citizens to believe this narrative.

This belief in gold as a safe haven, which is frequently supported by anecdotal evidence and personal experiences, is strengthened by a selective interpretation of history. Narrative 4: Historical Revisionism enhances the gold standard's appeal by giving a romanticized view of the past, stressing the perceived stability of the gold standard era while downplaying its disadvantages.

Narrative 4 Historical Revisionism

Figure 5 illustrates the evolution and revisionist views of the U.S. gold standard, highlighting key events like its establishment, modifications, and abandonment, along with renewed interest during periods of economic instability. This supports the finding that historical revisionism plays a role in the gold standard's appeal.

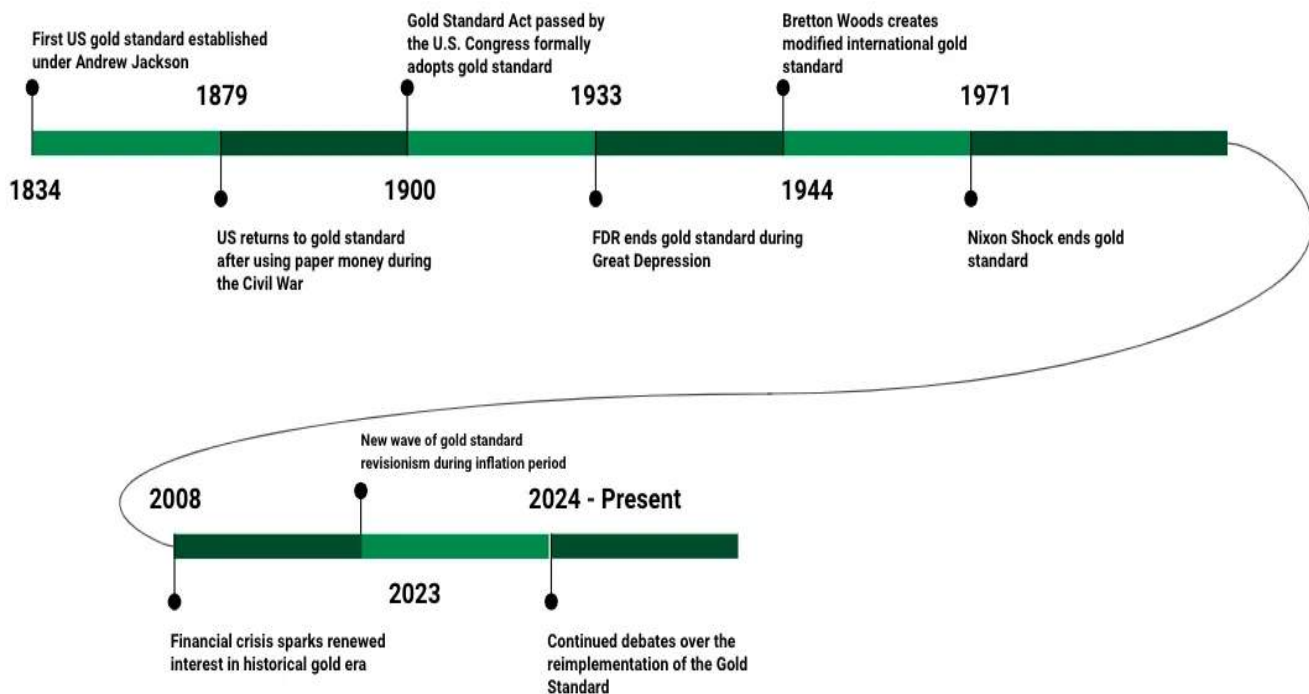


Figure 5: Evolution and Revisionist Views of the U.S. Gold Standard

The Historical Revisionism narrative suggests that economic conditions were more stable and prosperous when the gold standard was in place, particularly during the 19th and early 20th centuries (Domitrovic, 2019). This perspective asserts that the gold standard limited inflation, prevented reckless government spending, and contributed to long periods of economic growth and stability (Schiff, n.d.). Critics of modern monetary systems claim that moving away from the gold standard has led to increased economic volatility, inflation, and financial crises, portraying the era of the gold standard as a "golden age" of economic management and fiscal discipline (Forbes, 2024). However, this narrative often disregards the frequent deflations, financial panics, and economic sufferings that also marked the gold standard era, such as the Panic of 1907 and the Great Depression. The inherent limits of attaching currency to a finite resource are frequently overlooked, as is the gold standard's role to worsen economic downturns. This revisionist perspective's appeal stems in part from its simplicity and emotional relevance. By offering a nostalgic vision of an apparently simpler past, it provides comfort and a sense of security in the face of complex modern economic issues. Prominent figures like Jim Grant, Steve Forbes and Peter Schiff, along with think tanks such as the Mises Institute and AIER amplify this perspective by leveraging their platforms to argue that times under a gold standard were better

for everyday people. They make complex economic issues relatable, helping to sway public opinion by presenting the gold standard as a simple solution to modern financial problems.

These four narratives, each appealing to different aspects of public concern and amplified by various actors, collectively contribute to the persistent allure of the gold standard. Understanding the interplay of these narratives is critical for analyzing the larger debate about monetary policy and its impact on public perception. This leads to a consideration of the broader goals and implications of these narratives in the context of current economic discussions.

VI Discussion

Crucial Aims of the Narratives

These narratives' main goal is to advance the idea that restoring the gold standard would enhance monetary stability, control inflation, and restrict government involvement in the economy (Winfrey, 2024). A gold-backed currency, according to its proponents, would protect property rights and individual economic freedom by preventing governments from controlling the money supply. Economically unstable times like the 2008 financial crisis, the COVID-19 pandemic, and the inflationary worries of the early 2020s have given this narrative new life (Konkel, 2018). During periods of economic uncertainty, these narratives gain particular resonance by tapping into deep-seated fears about financial security and loss of control, with proponents skillfully connecting everyday experiences of rising prices and market volatility to broader arguments about monetary policy failure. The goal of presenting the gold standard as a solution for contemporary economic problems is to gain support for laws such as the Gold Standard Restoration Act (Mooney, 2023).

Significant challenges still exist, though, including political opposition from central bankers and policymakers who argue that a gold standard would reduce economic flexibility in times of crisis. Critics also draw attention to the difficulties in implementing a gold standard in a complicated global economy, the necessity of international cooperation, and the possibility that changes in the price of gold could lead to new economic instability. Attempts to reinstate the gold standard are made more difficult by public skepticism and the practical challenges of overseeing such a shift. Even though supporters continue to push for its adoption as a solution to contemporary economic issues, these difficulties highlight the issue's complexity.

Key Actors and Their Motivations

These narratives' proponents include both individuals and organizations, who have shared goals based on libertarian and conservative economic beliefs. Their objectives are to challenge the existing fiat money system, advocate for less government involvement in the economy, and eventually influence public opinion to support free-market principles (McMaken, 2024). They believe that bringing back the gold standard will improve fiscal discipline, promote individual liberty, and safeguard against alleged overreach by the government.

These narratives are being driven by key figures such as Ron Paul (McMaken, 2024), a former US congressman and prominent libertarian voice; Murray Rothbard (Rothbard, 2024), a

founding figure in the Austrian School of economics; and Steve Forbes (Forbes, 2024), editor-in-chief of Forbes Magazine and a vocal advocate for economic freedom and the gold standard. These individuals, along with others such as Lewis Lehrman and Ludwig von Mises (Murphy, 2024; Rothbard, 2024; Hughes, 2024), contribute to a body of literature that praises the perceived merits of a gold-backed currency while opposing fiat money and centralized economic control. Their influence is strengthened by groups such as the Heritage Foundation (Antoni, 2023; Fuelner, n.d.), the Mises Institute (Postel, 2013), and the CFA Institute (Hughes, 2024), which provide research and commentary on free-market economics and limited government. Initiatives like Project 2025 (Winfrey, 2024), which focuses on conservative policy change, also contribute to the current campaign for a gold standard. These players frequently condense complex economic theories into appealing narratives that resonate with the public, especially during periods of economic turmoil. Beyond policy influence, they also encourage investment in gold as a hedge against perceived economic risks, further reinforcing its perceived value and promoting financial independence (Kiyosaki, 2024). This intertwining of ideological advocacy and investment promotion creates a powerful force in shaping public discourse around monetary policy.

Historical Revisionism and Its Role

Historical revisionism plays a significant role in sustaining the appeal of the gold standard. Advocates often reinterpret the gold standard era as a "golden age" of economic stability and prosperity, downplaying its flaws and the economic crises it exacerbated (Postel, 2013). For example, revisionist accounts often minimize the deflationary periods, banking panics, and the gold standard's role in deepening the Great Depression, while ignoring the benefits of modern monetary policy flexibility. To change the public's perception of monetary history and encourage a return to gold-based systems, proponents of the gold standard try to disprove the widely held belief that it is outdated or irrelevant. Due to its apparent simplicity in addressing intricate economic issues, this narrative has gained popularity during times of financial instability, including the Great Depression, the inflation crisis of the 1970s, and the financial collapse of 2008 (Hughes, 2024; Maley, 2023).

Significance and Implications

The persistence of these narratives highlights the power of narratives in shaping economic beliefs and behaviors. Despite overwhelming academic consensus against the feasibility of the gold standard, these narratives continue to resonate because they address fundamental concerns about inflation, government control, and financial security. The impact of these narratives on policy decisions is seen by recent legislative proposals like the Gold Standard Restoration Act (Mooney, 2023). Furthermore, even if those decisions aren't always in line with sound economic principles, these narratives have the power to influence public financial habits and may influence investment decisions toward gold during uncertain economic times. In addition, digital platforms like Twitter (X), TikTok, and YouTube can amplify these narratives.

Influencers and commentators promote gold as a hedge against inflation or government overreach, often using memes, viral videos, or conspiratorial frameworks. This underscores the importance of understanding the ideological and cultural drivers behind economic policy debates, as they often influence public opinion and policy outcomes more than empirical evidence.

Hypothesis and Consistency with Other Studies

The study's findings support the hypothesis that the appeal of the gold standard is primarily driven by ideological and cultural narratives rather than by economic practicality. This aligns with Shiller's claim that narratives, frequently regardless of their accuracy, flourish on emotional resonance, identification, and shared values. These emotional and intellectual currents are specifically touched upon by the four narratives: historical revisionism, libertarian/conservative ideology, the critique of fiat money, and gold as a store of value. They speak to concerns about financial instability and governmental control by providing straightforward explanations and answers to intricate economic issues. These narratives, as Shiller's theory suggests, spread through contagion, particularly amplified by influential figures and media outlets, creating a sense of shared belief and identity among proponents. The results also correspond with historical analyses of the gold standard, which indicate that its popularity tends to increase during times of economic uncertainty, such as the Great Depression, the 2008 financial crisis, and the COVID-19 pandemic. Furthermore, the impact of influential people and groups in spreading these narratives is consistent with research on the bandwagon effect and narrative contagion, especially as it relates to cryptocurrency advocacy. The parallels between the gold standard narrative and the cryptocurrency boom, both fueled by celebrity endorsements and public skepticism of central banks, further underscore the power of narrative in driving economic behavior, as highlighted by Shiller.

Limitations and Future Research

This study has several shortcomings even though it offers insightful information about the myths surrounding the gold standard. The results rely on qualitative content analysis, which makes them interpretive and may not accurately capture the complexity of public opinion. Additionally, the study primarily looks at American, as well as pro-gold narratives and may not sufficiently account for perspectives from other nations. This work could be expanded upon in future research by employing quantitative methods, such as surveys or experiments, to gauge how these narratives affect public opinion and investing behavior. For example, surveys could measure the prevalence of belief in specific gold standard narratives and correlate these beliefs with investment choices and voting preferences. Large datasets of news stories and social media posts could be subjected to public sentiment analysis algorithms to measure the frequency and emotional impact of these narratives over time and correlate them with economic indicators such as inflation and gold prices. The relationship between economic sentiment and belief in gold standard narratives could be examined using datasets such as the University of Michigan

Consumer Sentiment Index or Gallup polls (University of Michigan: Consumer Sentiment, 2024, Gallup, Inc., 2024). In simulated market settings, experiments could examine the effects of exposure to various narratives on investment choices. Comparative research could also look at how these stories change depending on the political and cultural setting.

VII Conclusion

This study highlights four important narratives that maintain interest in the gold standard: criticism of fiat money and the Federal Reserve, libertarian and conservative ideology, belief in gold as a hedge against uncertainty, and historical revisionism. These narratives have real-world consequences, undermining public trust in monetary institutions and fueling trends such as the Gold Standard Restoration Act, and increased gold investments during economic downturns. These ideas, amplified by social media, present gold as a solution to government overreach or economic collapse. Understanding these narratives is critical for addressing the anxieties that underpin them and restoring trust in financial systems. These narratives continue, despite economists generally agreeing that returning to the gold standard would be damaging. Using Shiller's narrative economics to analyze these narratives, this study adds to existing literature by offering a framework for understanding how particular narratives maintain interest in economically outdated policies. It goes beyond merely refuting the gold standard's economic viability to explore the deeper factors that underlie its enduring appeal. Along with highlighting their impact on public discourse and their influence on policy decisions, we also identify the key individuals and organizations in charge of spreading these narratives. Furthermore, by linking these narratives to concrete policy proposals and observed financial behaviors, the study demonstrates their real-world significance.

The sustained popularity of the gold standard demonstrates significant patterns in how economic ideas spread. These narratives gain popularity during economic downturns and spread when propagated by prominent individuals and institutions. They appeal because they address fundamental worries about government control, economic stability, and financial security.

Understanding these narratives is critical for policymakers in communicating monetary policy and addressing public concerns about financial stability. For instance, through accessible public education programs, central banks could increase transparency by outlining their policy objectives and the reasoning behind their choices. Examining different monetary policy frameworks, including nominal GDP targeting, may provide a more direct response to worries about inflation and economic volatility. The appeal of revisionist narratives may also be lessened by increasing public awareness of the gold standard's shortcomings and historical shortcomings. The persistence of gold standard advocacy suggests that central banks should better address public concerns about monetary policy.

Looking ahead, while the gold standard may not be suitable for modern monetary policy, the narratives that sustain it reveal legitimate public worries about monetary stability, inflation, and political power. Rather than ignoring gold standard advocacy, addressing these fundamental concerns could result in more inclusive and transparent economic policy. Policymakers can

increase public trust and understanding by addressing these issues, such as through improved communication from central banks, policies that improve financial stability, and open discussion regarding monetary system trade-offs. This strategy not only bridges gaps, but also opens the door to more constructive and informed talks about the future of economic policy.

Works Cited

- Antoni, E. (2023). Time to end the Fed and its mismanagement of our economy. *The Heritage Foundation*. <https://www.heritage.org>
- Bromberg, M. (2024, June 10). 7 good reasons to own gold. *Investopedia*. <https://www.investopedia.com>
- Buckfire, H. (2024, July 18). Would project 2025 abolish the Federal Reserve? What about reinstituting the gold standard? - SPDR gold trust (ARCA:GLD). *Benzinga*. <https://www.benzinga.com>
- Cooper, R. N., Dornbusch, R., & Hall, R. E. (1982). The gold standard: Historical facts and future prospects. *Brookings Papers on Economic Activity*, 13(2), 361–383. <https://doi.org>
- Copeland, C. (2024, August 5). Can we go back to the gold standard? Project 2025 thinks so. *Fort Worth Star-Telegram*. <https://www.star-telegram.com>
- Domitrovic, B. (2019, July 16). The argument for returning to the gold standard: The lessons of contemporary history. *Hillsdale College*. <https://www.hillsdale.edu>
- Duggan, W. (2024, May 1). Is gold an inflation hedge? *Forbes*. <https://www.forbes.com>
- Fallert, N. (2021, April 16). Everything Elon Musk has said about Dogecoin. *Newsweek*. <https://www.newsweek.com>
- Feulner, E. J. (n.d.). Towards a stable monetary policy: Monetarism vs. the gold standard. *The Heritage Foundation*. <https://www.heritage.org>
- Five reasons why gold remains the ultimate safe haven asset. (n.d.). *The Royal Mint*. <https://www.royalmint.com>
- Forbes, S. (2024a, April 25). Believe it or not, the world is beginning to lurch forward to a new gold standard. *YouTube*. <https://www.youtube.com>
- Forbes, S. (2024c, June 4). The signs are there: The gold standard is coming back. *Forbes*. <https://www.forbes.com>
- Gallup, Inc. (2024, November 26). *Economic Confidence | Gallup Topic*. <https://news.gallup.com/topic/economic-confidence.aspx>
- Gold price. (2002, September 20). *Gold Price*. <https://goldprice.org>
- Gold: The timeless store of value: Exploring the role of gold. (2024, June 18). *FasterCapital*. <https://fastercapital.com>
- Hughes, B. (2023, October 19). Markets in chaos: A return to the gold standard? *CFA Institute Enterprising Investor*. <https://blogs.cfainstitute.org>
- Kiyosaki, R. (2024, June 26). The future of the gold standard and the collapse of the dollar. *Rich Dad*. <https://www.richdad.com>
- Konkel, L. (2018, December 7). How did the gold standard contribute to the Great Depression? *History.com*. <https://www.history.com>
- Konkel, L. (2023, December 7). How did the gold standard contribute to the Great Depression? *History.com*. <https://www.history.com>
- Lewis, N. (2015, December 10). 40 out of 40 economists agree: The monetary system that made America great is nonsense. *Forbes*. <https://www.forbes.com>

- Lioudis, N. (n.d.). What is the gold standard? History and collapse. *Investopedia*.
<https://www.investopedia.com>
- Maley, K. (2023, October 31). Why Jim Grant believes gold will have its day. *Australian Financial Review*. <https://www.afr.com>
- McMaken, R. (2024, April 11). Alan Greenspan admits Ron Paul was right about gold. *Mises Institute*. <https://mises.org>
- Mooney, A. X. (2023, March 30). Gold standard restoration act 118th Congress (2023-2024). *Congress.gov*. <https://www.congress.gov>
- Mueller, J. D. (2015, January 15). Where the gold standard act of 1984 came from—and why it’s still relevant. *Ethics & Public Policy Center*. <https://eppc.org>
- Murphy, R. P. (2024, April 11). The gold standard: Myths and lies. *Mises Institute*.
<https://mises.org>
- Narrative (n.). (2024, March). *Oxford English Dictionary*. <https://www.oed.com>
- O’Brien, M. (2012, August 26). Why the gold standard is the world’s worst economic idea, in 2 charts. *The Atlantic*. <https://www.theatlantic.com>
- Paul, B. (2024, April 30). Gold standard: Definition, evolution, and applications. *SuperMoney*.
<https://www.supermoney.com>
- Postel, C. (2013, September 17). Why conservatives spin fairytales about the gold standard. *Reuters*. <https://www.reuters.com>
- Rothbard, M. N. (2024, April 11). The case for a 100 percent gold dollar. *Mises Institute*.
<https://mises.org>
- Schiff, P. (n.d.). Peter Schiff: The gold standard didn’t fail. *SchiffGold*.
<https://www.schiffgold.com>
- Schoenholtz, S. C., & K. (2017, January 6). Why a gold standard is a very bad idea. *Money, Banking and Financial Markets*. <https://www.moneyandbanking.com>
- Scott, R. (2023, March 21). Federal Reserve out of control: Americans shouldn’t have to pay for Wall St-Washington mistakes. *U.S. Senator Rick Scott*. <https://www.rickscott.senate.gov>
- Shiller, R. J. (2019). *Narrative economics: How stories go viral & drive major economic events, with a new preface by the author*. Princeton University Press.
- University of Michigan: Consumer sentiment*. (2024, December 20).
<https://fred.stlouisfed.org/series/UMCSENT>
- Vandenboss, K. (2024, March 20). “The dollar’s days as the reserve currency are numbered” - Peter Schiff says central banks are “preparing for a new gold standard.” *Yahoo! Finance*.
<https://finance.yahoo.com>
- Winfrey, P. (2024). Federal Reserve. *Project 2025*. <https://static.project2025.org>

How do Social Norms and Peer Effects Affect Household Debt Accumulation and Financial Stability? A review study By Shraddha Balaji

Abstract

The impact of peer pressure and social norms on household debt accumulation and financial stability is reviewed in this paper. According to earlier studies reported in the literature, both social norms and peer pressure influence debt levels. The social comparisons within the cultures normalize borrowing, leading to greater debt levels. Additionally, in social situations, peer influences are significant. People usually tend to borrow or save money in similar ways as their peers. Debt is seen to be highly common in households in which peers are expecting a larger amount of income. To create effective financial policies, understanding these behavioral tendencies is required. The impact of peer pressure can be reduced by encouraging financial responsibility and improving general economic stability by modifying societal norms surrounding borrowing and saving through awareness-building and educational initiatives. To successfully do so, the policies should address both the social and behavioral aspects of debt accumulation.

Keywords

social norms, peer effects, household debt, financial stability, financial policies

I. INTRODUCTION

The paper reviews the reports and publications of previous research to analyze the relationship between social standards, peer pressure, and household debt accumulation. The household debt level determines the ability to maintain financial stability in a household and is shaped by social norms and peer influences. The social norms are behavioral expectations of the individuals within a social group, and individuals align their behavior to fit in their respective social circle. Borrowing to live life as expected in your social norm is one such behavior. Reports have shown that debt levels rise proportionally when borrowing becomes more common in a community. The effects of peer group thinking and comparisons within the peer group also affect the debt level, as people often borrow or save in similar ways as their friends or peers. The review sheds light on the impact of social norms on household debt levels and how effective financial policy decisions may foster a change in this pattern. Through social education and marketing strategies, the social attitude toward borrowing and saving can be reshaped, and the impact of peer pressure can be reduced, leading to conscious financial responsibility among the members of society, thus improving economic stability. The analysis emphasizes the importance of policies addressing the behavioral and societal factors contributing to debt accumulation.

The primary research question in this paper is, Do social norms and peer effects affect household debt accumulation and financial stability? The hypothesis is that social norms (that normalize borrowing and peer influences that encourage mimicking of consumption behaviors) contribute significantly to higher levels of household debt, which in turn affects financial

stability. The aim of this research was to explore the extent to which social norms and peer effects drive household debt accumulation and to assess how these factors impact the broader financial stability of households and the economy. By understanding these dynamics, the research seeks to provide insights that could help in formulating more effective financial policies and educational programs. This paper attempts to expand on these studies by analyzing how these social factors influence debt accumulation, debt measures, and consequently financial security, filling the gap regarding the macroeconomic impact of these actions.

METHODOLOGY: Relating to the domain of peer influence and financial stability, the researcher undertook this search into peer influence literature in a very systematic manner. The use of keywords “**peer effects**,” “**social norms**,” “**financial stability**,” “**debt accumulation**,” and “**household savings**” helped to identify relevant papers using databases. The further narrowing of the investigation’s scope was done by including peer-reviewed, working, and institutional publications in the searches; hence, high-level sources were incorporated only. The next step involved allocating the papers according to the research question and the most significant themes. The studies falling under this category were organized into such subtopics as social comparison theory and peer effects in financial choices as well as debt and saving behavior of people with relation to social norms. In this process, I synthesized results from the various papers to answer the central question of the effects of peer influence on financial stability and the possible synergies and divergence that characterize the available literature on the topic. This introduction outlines the approach followed in collecting and synthesizing papers with regards to the literature review. The discussion of specific studies and their analysis will be covered in the main body of the Paper ‘Literature Review’ section in due course.

II. LITERATURE REVIEW

Research on the influence of social norms and peer effects on borrowing in general and household debt in particular has received some attention. Lusardi & Tufano (2015)¹ determined that those communities with normalization of borrowing practices are more likely to see excessive debt being accumulated. Also, Bertrand & Morse (2016)² showed that individuals take on enormous debt levels due to the social notion that people themselves rationalize such debt levels based on those around them. Social comparison theory³ is another theoretical framework that posits that individuals engage in certain financial activities due to the impact of their peers. This literature highlights that social and peer factors impact financial decisions, hence the need for policies geared toward achieving financial stability that incorporate such factors.

Social norms and consumption

The importance of peer income for consumption was first formalized in Duesenberry’s 1949 relative income hypothesis⁴. According to this hypothesis, households with incomes below average in their social circle tend to consume a larger share of their income to keep up with peers. Kuhn et al. (2011)⁵ found that exogenous variations in income from winning a Dutch

postal lottery tend to influence not only the durables purchased by winners but also the probability that neighbors will buy a new car. The financing of the neighbors' induced spending is not explored. The empirical literature indeed suggests that peer effects can influence household consumption^{2, 5, 6, 7}, stock market participation⁸, housing, mortgage leverage choice^{9, 10} and personal bankruptcy risk^{11, 12}. Survey evidence in Georgarakos, Haliassos, and Pasini (2014)¹³ shows that the perceived income rank relative to peers' affects borrowing. Kuhn et al. (2011)⁵ use data from the Dutch Postcode Lottery and survey data on consumption to study how income shocks affect winners' and their neighbors' consumption and happiness. A key finding is that neighbors of winners increase consumption and are more likely to own a new car in the years after their neighbor wins in a lottery. As pointed out by Georgarakos et al. (2014)¹³, a peer effect that affects debt need not reflect a peer effect through consumption. Rayo and Becker (2006)¹⁴ provide a model with one simple mechanism linking conspicuous consumption and borrowing. Social status is linked to visible goods, which are also costly, durable goods. Thus, for economic agents that want to smooth consumption, status-driven consumption leads to more borrowing or less saving. A long strand of empirical literature has sought to find evidence of social image as a determinant of consumption, particularly visible consumption.

Bertrand and Morse (2016)² find evidence of "trickle-down consumption," i.e., that poorer households spend more on visible goods if exposed to higher top income levels, with the implication that they save less than comparable households in other regions do. Aiming at understanding mechanisms driving peer effects, Bursztyn, Ederer, Ferman, and Yuchtman (2014)¹⁵ conduct a field experiment and find evidence that both social learning (i.e., learning about the value of an asset through peers' purchases of the asset) and social utility (i.e., the utility from owning an asset increases with peers' possession of the same asset) affect investment decisions. Finally, with identification through "friends-of-friends" networks, De Giorgi et al. (2020)⁷ build on work by Bramoullé, Djebbari, and Fortin (2009)¹⁶ and De Giorgi, Pellizzari, and Redaelli (2010)¹⁷ to study consumption network effects. With Danish household-level data and household members' workplaces as the social network, they find small but significant network effects in consumption and show that their implied government spending multiplier depends on the policy's targeted section (poor or rich).

Data on Debt

A. Abdul Jamal et al. (2015)¹⁸ reported in Malaysia that the sharp increase in the ratio of household debt to disposable income, particularly over the past ten years, has raised doubts on the ability of Malaysians to save and plan for their future. As a result, total household savings remained low, and there is a high risk that Malaysians would not have sufficient savings for their retirement. Even more worrying is that young adults reportedly are the main group who are trapped in this financial difficulty. The issue has raised concern about the need to educate young Malaysian adults on the fundamental importance of savings in order to ensure financial sufficiency for their retirement. With regards to savings behavior, the needs for savings are different among individuals due to different mindsets, behaviors, knowledge, and social

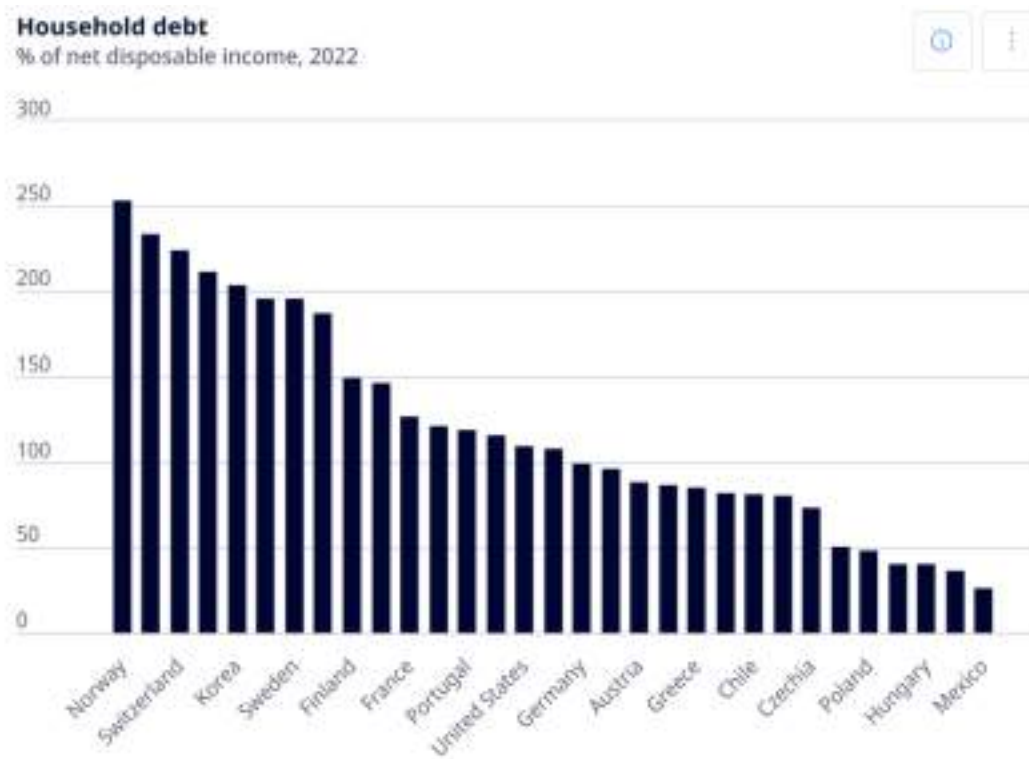
environments. In some cultures, taking on debt may be seen as a sign of financial irresponsibility, discouraging borrowing. Conversely, in other contexts, debt might be normalized or even encouraged, leading to higher levels of borrowing. Societal attitudes towards saving can also impact debt levels. In societies where saving is valued, households might be more cautious about accumulating debt.

Social norms and borrowing

Almenberg, Lusardi, Säve-Söderbergh & Vestman (2020)¹⁹ suggested that people may decide to save, or may refrain from borrowing, because they have internalized a social norm that discourages debt. Social norms that discourage borrowing are not by definition good or bad; for example, a social norm against debt that deters talented people from investing in human capital reduces economic efficiency. And a social norm against debt may inhibit consumption smoothing in a way that is similar to a liquidity constraint, with comparable welfare losses. Social norms against debt could also explain why many households pay a sizable premium for less salient debt.

III. HOUSEHOLD DEBT

Figure 1. Household debt across countries in 2022



High household debt levels, as shown in certain countries' patterns (FIG.1²⁰), follow the uptake of financial products focusing on borrowing, becoming the norm. In cultures where one has to bury oneself in all manner of debt to achieve some level of social status, one may be socialized into incurring debts that one cannot afford. This corresponds to economic research

that proves that common moral views and pressure from other individuals lead to more debt. The differences in household debt levels across countries also suggest that cultural factors related to socialization or even normative behavior may be important predictors of borrowing behavior in subpopulations within countries. In countries where the ratio of household debt to household income is relatively low, it is likely that there is an adverse social attitude towards excessive personal debt or that there is strong personal financial discipline. Furthermore, in countries with higher debt ratios, there may be cultures of relative and/or absolute poverty that glorify borrowing.

High household debt levels in countries can also pose some risks for financial instability, especially when there is a large part of the households that borrow more than they can afford. There is likely to be a more serious knock on the economy when these debts look like they will not be paid back, as collective debt levels can lead to recession. Fig. 1 also highlights the requirement for appropriate intervention strategies so as to tackle the social and psychological aspects behind them, which make people borrow money. The findings support the idea that using peer programs or networks to inculcate financial education that shapes taste can help change the stigma behind debt.

Peer Effects and Financial Stability

Bursztyn et al. (2014)¹⁵: The findings of this research reveal that peer influences can influence financial decision-making, depending on the prevailing social norms. One subject who shared the results of an experiment claimed, for instance, that people are more inclined to buy a financial asset if they know that a peer bought such an asset (social learning). The study also found that knowing that a peer owned an asset increased the probability of an individual purchasing it even more ('social utility'). Social utility was further claimed to be motivated by the want to keep up with the financial level of other people. The study also indicates that some people considered it important for their decision and that possessing the asset would allow them to discuss it with their peers. In societies with high-borrowing cultures, individuals push one another into more debts due to the tendencies of taking up similar financial activities. On the other hand, in societies that foster saving, individuals are helped by the community to save more by lessening the amount of debt incurred through peer effects. This dual nature of peer effects strengthens the need to incorporate social dimensions when analyzing financial activities.

Based on the research carried out by Beshears et al. (2011)²¹ and Karlan et al. (2014)²², it is clear that policies that seek to enhance the financial wellbeing of people should not be only focused on education. They need to play an active role in transforming the social aspects and influencing positive aspects of individuals within their peers. Changes of social norms that discourage consumer debt and other pressures to prove one's financial worth to others bearing possibly negative consequences can be achieved through sound financial education programs that highlight the consequences of debt and highlight the positives of healthy finance management practices.

Financial Stability

Education on financial stability is crucial not just for the individual but also for the community's economic well-being. It covers a lot, from one's income to savings and investment strategies. It, however, goes beyond strategies and structures as it positively influences a person's behavioral patterns. Obtaining knowledge about these behavioral patterns assists policymakers, financial institutions, and individuals in making optimum financial choices. The objective of this literature review is to find out what theory-based behavioral patterns exist that may be helpful for assessing a country's level of financial stability from the available literature to date and what the economic consequences of these patterns are.

Social networks involving peer groups in the evaluation of financial risk exhibit certain features typical of different branches of science, such as psychology and economics. One of the most popular explanations of self-evaluation is known as the Social Comparison Theory³, which describes how one compares opinions and abilities with other individuals in order to assess the status of oneself. In this case, applying Social Comparison Theory within the financial domain implies that people balance various financial activities, such as savings or borrowing, against how well similar actions are carried out by other people.

Peer effects can also be explained by theories from behavioral economics. The idea of bounded rationality when making decisions, which was first proposed in Simon (1955)²³, implies that due to limited information and other cognitive limitations, people often make use of techniques and social norms available in their environment. The perspective of the social environment in the financial area, normally, means the same as its peers in that sense, and as observed, one engages in herd activities, usually to the worse of the individual's financial health.

Likewise, most of the focus in the studies has been on borrowing behavior patterns; however, peer influence is significant too in savings patterns. Gale and Scholz (1994)²⁴ estimate that increasing the contribution limits for IRAs would have little effect on the overall rates of saving. They remarked that people were more inclined to save when others, especially friends, were more concerned about saving than spending. This means that if saving is regarded as the norm amongst a certain populace, it is possible that peer effects can be imposed on individuals and thus promote the building of financial cushions for insurances against AS shocks.

In the same way, other research has analyzed the impact of social norms on individuals' decisions to join pension schemes. Their investigation showed that the likelihood of employees contributing to saving plans increased with the participation of their coworkers, showing peer influences in making long-term financial decisions. This particular knowledge is useful within the scope of behavioral change strategies that seek to increase saving levels as it provides that peer support can be purposefully utilized to achieve improved outcomes.

Behavioral Economics and Financial Stability

Behavioral economics is an area of study that is concerned with how economics and psychology can be integrated to help interpret situations where humans make decisions contrary to the traditional economic model of behavior. The standard economic theories hold that the

maximization of utility given the rations is the rational behavior of humans. In contrast, behavioral economics seeks to highlight the fact that human beings are susceptible to erroneous decisions, mainly through cognitive biases, emotions, or social pressures, which usually could affect financial stability.

Prospect theory, developed by Kahneman and Tversky in 1979²⁵, is one of the main theories of behavioral economics. This theory states that people will have in relation to gains and in relation to losses some attitude and behavior known to be economically rational and risk avoidant in nature in case gains are involved and risk seeking when it comes to experiencing losses. This distortion in how people perceive risks can lead to poor choices in terms of financial options, such as taking too much time to exit an unprofitable position or exiting a winning position too soon.

Behavioral patterns are very essential in promoting financial stability. Understanding these behaviors, however, enables policymakers and financial institutions to tailor out interventions that would help promote better financial decision-making.

Improving financial literacy and education of the public are the most essential means of reducing behavioral effects that potentially jeopardize financial stability. Besides basic literacy and education, behavioral-based interventions, including nudges and the use of commitment devices, are likely flawless methods of enhancing financial literacy. Here, more studies may be called for to elaborate on some of these interventions and their relevance to people's backgrounds.

According to Lusardi and Mitchell (2014)²⁶, finance literacy affects one's tendency to engage in the anticipated financial behavior, and the authors demonstrate that those who are more financially educated are more likely to save for retirement and are less likely to get into high debt. For the opposite reasons, those who are poor at finance management are more deeply in debt and make more financial mistakes.

IV. DISCUSSION

The outcomes of the literature review based on the hypothesis suggested that social norms that lend credence to borrowing lead the individual borrower to engage in self-negotiation of their debt levels when placed among their peers. One of the most important unfinished propositions is how peer influence drives people to borrow more money. For example, once a few of their peers increase their debt, individuals are much more willing to follow suit. Instead, debt accumulation can be countered through the use of specific strategies that target changing social norms and enhancing positive social externalities.

From the findings of this research, it is evident that social norms and peer effects have a significant role in affecting the amount of debt households are willing to self-accumulate. These social factors can in one way or another affect financial stability.

Analysis of Social Norms and Debt Accumulation

As further evidence, it is challenging to highlight any noticeable difference between these

two studies with regards to sociological factors underlying debt. In the case of domestic borrowing, particularly from friends and family, again it is well known that some say you can never borrow money to pay your debts or it is forbidden to ever borrow money from a relative. This is a clear case of overriding personal economic sense based on social norms.

In borrowing cultures, the population is inclined to incur debts, usually due to social forces and a wish to keep up with the wealthier members of society. This backs up the assertion that social influences are indeed one of the reasons for increased household levels of debt in the populations. However, there is an important caveat: social norms are dual in nature. While in some cultures the physical borrowing behavior is normalized and the community accumulates mountains of debts, in other communities that have a negative mile in debt societies, the practices lead to moderation of the financial risks. This shows that social norms at one point can become a risk factor and at the other end, a risk prevention factor, depending on that culture. This is consistent with previous studies highlighting the individual variation of social norms across cultures and the different ways these norms manifest into financial actions.

V. CONCLUSION

In this work, it has been proven and investigated that social norms and peer effects are at play in regard to household debt levels and financial wellbeing. In societies where taking loans is commonplace, people are bound to contract unwarranted debts propelled by the need to keep pace with the spending habits of the rich. This in comparison is opposite in societies that encourage saving in that the social influences in a community motivate members to spend wisely, hence resulting in low levels of debt and increased financial stability. These results emphasize the importance of not only the information aspect of the financial literacy programs and campaigns but also the behavioral aspect, which aims at changing negative norms and embracing positive influence from peer groups. This allows a more specific approach to the design of policy measures. Assimilating characteristics of processes like this in specific packages of policy interventions has positive effects on the functioning of society and every separate family. Further work should investigate how these social aspects function and whether they would be effective in different settings and what other moderators may be involved that may in turn affect behavior towards debt.

In order to continue working with these social dynamics', it would be necessary for policymakers and educators to devise measures that promote financial security and reduce the systemic risks related to household debts.

People frequently take after their peers when it comes to money management. People may be more inclined to take on debt for large purchases if they see their friends or neighbors doing so. Peer groups frequently share information about interest rates, credit availability, and financial products. This may result in better informed—or worse informed—decisions about borrowing. Social networks can have an impact on financial behavior as well. Peer borrowing habits and financial choices can create a feedback loop that either raises or lowers the amount of debt in the community.

Homes may get over leveraged when peer pressure and social norms encourage excessive borrowing. This may raise the chance of defaults, especially if the economy worsens and more widespread financial instability results. Asset price bubbles may occur if a large number of households simultaneously take on more debt in order to invest in assets like real estate (motivated by peer pressure or social norms). The economic collapse that follows the burst of these bubbles may jeopardize financial stability. The accumulation of individual household debt behaviors can put the financial system at risk on a systemic level due to peer pressure and social norms. A population's high levels of household debt can increase the economy's susceptibility to shocks.

Understanding these dynamics is crucial for policymakers aiming to promote financial stability and for individuals seeking to manage their personal finances responsibly.

VI. Limitations and Future Research

Despite the usefulness of this study, it is also wise to accept shortcomings that are available in this report. One considerable limitation is the difference in peer pressures or social norms or how these factors play out in various cultures and socioeconomic environments. Most of the studies that were reviewed in this paper were conducted in particular regions within countries, which may not represent the scope of peer relationships and social norms as experienced around the globe. In order to understand the viability of the results, it is recommended that further research combine ethnographies with such social dynamics as those studied in the present paper.

This poses another limitation since there are many other unaccounted variables, such as personality traits of the individual or certain economic circumstances, that might affect the way social norms, peer influence, or behavioral financial work. While the focus of this research was on only social factors, expansion of this study to consider these other factors may improve understanding of behavioral finance microscale. If appropriate measures are taken to address these shortcomings, then the findings of the research will have positive implications in relation to what other researchers are likely to do in the same field. If appropriate measures are taken to address these shortcomings, then the findings of the research will have positive implications in relation to what other researchers are likely to do in the same field.

Acknowledgements

I acknowledge the guidance of my mentor, Philip Liang from the University of Chicago, in the development of this research paper. I appreciate the encouragement and efforts of my writing coach, Maria Noelia Romero.

Works Cited

1. A. Lusardi & P. Tufano. Debt literacy, financial experiences, and overindebtedness. *Journal of Pension Economics and Finance*, **14**, 329–365. (2015).
2. M. Bertrand & A. Morse. Trickle-down consumption. *The Review of Economics and Statistics*, **98**, 863–879. (2016).
3. L. Festinger. A theory of social comparison processes. *Human Relations*, **7**, 117–140. (1954).
4. R. Adhikari. (2022). Prevalence of Duesenberry’s relative income hypothesis in Nepalese economy. *International Review of Business and Economics*, **7**, Iss. 1, Article 2. (2022). DOI <https://doi.org/10.56902/IRBE.2022.7.1.2>. Available at: <https://digitalcommons.du.edu/irbe/vol7/iss1/2>
5. P. Kuhn, P. Kooreman, A. Soetevent, and A. Kapteyn. The effects of lottery prizes on winners and their neighbors: Evidence from the Dutch postcode lottery. *American Economic Review*, **101**, 2226–2247. (2011).
6. S. Agarwal, W. Qian, and X. Zou. Thy neighbor’s misfortune: peer effect on consumption. *American Economic Journal: Economic Policy*, **13**, 1–25. (2021).
7. G. De Giorgi, A. Frederiksen, and L. Pistaferri. Consumption network effects. *The Review of Economic Studies*, **87**, 130–163. (2020).
8. M. Kaustia and S. Knüpfer. Peer performance and stock market entry. *Journal of Financial Economics*, **104**, 321–338. (2012).
9. B. Han, D. Hirshleifer, and J. Walden. Visibility bias in the transmission of consumption beliefs and undersaving. *The Journal of Finance*, **78**, 1647–1704. (2023).
10. M. Bailey, E. Dávila, T. Kuchler, and J. Stroebe. House price beliefs and mortgage leverage choice. *The Review of Economic Studies*, **86**, 2403–2452. (2019).
11. S. Agarwal, V. Mikhed, and B. Scholnick. Peers’ income and financial distress: Evidence from lottery winners and neighboring bankruptcies. *The Review of Financial Studies*, **33**, 433–572 (2020).
12. P. Roth. Inequality, relative deprivation, and financial distress—evidence from Swedish register data. (2020). Available at SSRN 3746651. <https://www.ifn.se/media/pxcla1yw/wp1374.pdf>
13. D. Georgarakos, M. Haliassos, and G. Pasini. Household debt and social interactions. *The Review of Financial Studies*, **27**, 1404–1433. (2014).
14. L. Rayo and G. S. Becker. Peer comparisons and consumer debt. *University of Chicago Law Review*, **73**, 231–248. (2006).
15. L. Bursztyn, F. Ederer, B. Ferman, & N. Yuchtman. Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, **82**, 1273–1301. (2014).
16. Y. Bramoullé, H. Djebbari, and B. Fortin. Identification of peer effects through social networks. *Journal of Econometrics*, **150**, 41–55. (2009).
17. G. De Giorgi, M. Pellizzari, and S. Redaelli. Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, **2**,

- 241–275. (2010).
18. A. A. Abdul Jamal, W. K. Ramlan, R. Mohidin, M. A. Karim, & Z. Osman. The effects of social influence and financial literacy on savings behavior: A study on students of higher learning institutions in Kota Kinabalu, Sabah. *International Journal of Business and Social Science*, **6**, 110–119. (2015).
 19. J. Almenberg, A. Lusardi, J. Säve-Söderbergh, & R. Vestman. Attitudes toward debt and debt behavior. *The Scandinavian Journal of Economics*, **123**, 780–809. (2020).
 20. Household debt. <https://www.oecd.org/en/data/indicators/household-debt.html>.
 21. J. Beshears, J. J. Choi, D. Laibson, & B. C. Madrian. Behavioral economics perspectives on public sector pension plans. *Journal of Pension Economics and Finance*, **10**, 315–336. (2011).
 22. D. Karlan, A. L. Ratan, & J. Zinman. Savings by and for the poor: A research review and agenda. *Review of Income and Wealth*, 60(1), 37–78. (2014).
 23. H. A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, **69**, 99–118. (1955).
 24. W. G. Gale & J. K. Scholz. Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives*, **8**, 55–80. (1994).
 25. D. Kahneman & A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263–291. (1979).
 26. A. Lusardi and O. S. Mitchell. The economic importance of financial literacy: theory and evidence. *Journal of Economic Literature*, **52**, 5–44. (2014).

Assessing the Efficiency of Data Collection: A Case Study for the Royal Bengal Tiger

By Virenpratap Singh Bhaika

Abstract

India has experienced significant success in its tiger conservation through initiatives like the government-led Project Tiger initiative. Since Project Tiger's inception in the 1970s, the initiative has been attributed to a 6.1% annual increase in the Indian tiger population. However, ongoing challenges for tiger conservation—such as habitat loss and human-wildlife conflict—remain prevalent. Altogether, this enduring problem for tiger conservation underscores the need for continuous and accurate monitoring methods in order to efficiently and correctly maintain updated counts on India's tiger population. To this end, I used agent-based modeling in NetLogo to compare the time efficiency of two widely used methods for tiger population estimation: camera trap surveys and vehicle surveys. Results demonstrate that camera traps require significantly less time to count tigers, thereby providing evidence that supports their use as a more efficient and consistent population monitoring tool.

Introduction

India is home to approximately 3,700 tigers (*Panthera tigris*) — one of the most influential and commonly occurring animals in Indian culture and mythology (Ministry of Environment, Forest and Climate Change, 2023). As the national animal of India, tigers are key features in national folklore and are prized as a national treasure. In addition to their societal value, tigers serve a key ecological role in the ecosystem functioning as an apex predator. However, despite their ecological and cultural significance, tiger populations are endangered across their range (Goodrich et al., 2022). Major threats to tigers include poaching, habitat loss, and human-wildlife conflict. India's rapid development has taken the form of highways, power plants, and expansion of human settlements. In turn, this depletes habitat size and continuity, harming a species that needs large territory for adequate resources (Sinha et al., 2012). These development projects not only encroach upon tiger habitat but also increasingly push tigers into contact with humans. This can result in the injury or death of humans and livestock (Chowdhury et al., 2016). Such conflict can lead to retaliatory killings and/or permanent translocation of tigers into holding facilities. Human-tiger conflict is particularly high in the national park buffer zones, areas which allow humans and tigers to share the same space.

Government agencies, NGOs, and community groups are three key pillars for tiger conservation in India. Non-government efforts are led by institutions like the Wildlife Institute of India (WII) and World Wide Fund for Nature (WWF). Their work targets research and community-based conservation education. Local community groups work exclusively for the benefit of the local context in which they are based, including tribes living in and around national parks and protected areas. They typically advocate for the voices of these communities to be heard, including for compensation for the lost livestock or human life. India's government efforts to support tiger conservation is exemplified by its flagship Project Tiger (PT) initiative which

works to restore and protect tigers and their habitat. PT was established in the 1970's as part of the then Indian Prime Minister (Indira Gandhi) plan to bring back India's tiger population which, at the time, was estimated around 1,800 individuals (Jhala et al., 2021). PT is a collective of many conservation efforts, including NGO and community-based work, from on-the-ground intervention to community education about conflict. The laws derived from Project Tiger include those that criminalised the poaching and illegal trade of tigers (Sethi, 2022). Other efforts include educational programmes for villages to learn about the ecological and tourism significance of tigers (Arjunan et al., 2006).

PT's efforts have been credited with India's substantial increase in its tiger population from about 1,800 in the 1970s to 3,700 today — a 6.1% annual increase since PT's inception (Ministry of Environment, Forest and Climate Change, 2023). Today, India hosts over 70% of the world's wild tiger population. With this, India succeeded in doubling its tiger population between 2014 and 2018. This benchmark was achieved before the 2022 deadline set by the St. Petersburg Declaration, which stated that all thirteen countries with wild tiger populations should double their numbers by 2022 (Diepstraten et al., 2022). However, questions have been raised about the choice of tiger count methodology and subsequent reporting of India's tiger population 'growth' has been met with criticism. Given that monitoring the tiger population is a laborious and financially intensive task for PT, the documented tiger population increase has also coincided with methodological and technological updates in wildlife monitoring. This has placed into question whether the population genuinely increased and/or whether tiger counting has simply become more accurate (Gopalaswamy et al., 2022). Ultimately, this discussion demonstrates the importance of accurate wildlife counts as well as transparent, understandable science communication.

In light of these challenges, tiger conservation efforts require efficient, accurate, and cost-effective data collection methods that can be scaled across the country's vast tiger habitats. Traditional methods such as vehicle surveys and camera traps have both been used to estimate tiger populations, but their effectiveness and time efficiency have been debated. Recent literature suggests that while camera traps offer more accurate population estimates, vehicle surveys are often used due to their accessibility and ability to monitor larger areas in real-time (Karanth et al., 2006; O'Connell et al., 2011). To address this gap, this study aimed to assess the time efficiency of these two methods by using agent-based modeling in NetLogo, a tool that allows for the simulation of various scenarios and provides quantifiable data for comparison. Below, I evaluate their relative time efficiencies against their overall utility in different tiger conservation scenarios.

Methods

I used NetLogo (Wilensky, 1999) to simulate two different methods for counting tigers: camera traps and vehicle surveys. NetLogo is a simulation tool for agent-based modeling, which allows for the exploration of different scenarios without costly or time-consuming fieldwork (Wilensky and Rand, 2015; Railsback and Grimm, 2019). All simulations were conducted in the

default environment of 1,089 patches. Thus, the study area was considered to be 1,089 km². Ten tigers were generated at the start of each simulation, representative of normal tiger densities (e.g., 0.1-20 individuals per 100 km²; Yumnam et al., 2014).

In the camera trap simulations (Figure 1A), nine stationary camera traps were placed in a symmetric grid and ten tigers were distributed randomly across the environment. When initiated, the tigers proceeded to move at random. A tiger became ‘counted’ when it shared a patch with a stationary camera – representative of a tiger being captured on camera in the wild. In the vehicle survey simulations (Figure 1B), tigers and vehicles were represented as moving agents within a closed environment. Initially, ten tigers and five vehicles were distributed randomly across the landscape and, when initiated, proceeded to move at random. A tiger became ‘counted’ when it shared a patch with a vehicle – representative of a tiger being visible to a surveyor in a vehicle, in real life. Both models stopped when all tigers were counted. Each simulation type was executed 50 times.

I recorded the time (ticks) taken to count all tigers and calculated the minimum, maximum, and mean (with standard deviation) time taken for vehicles and cameras to count all tigers. To compare the time required to count tigers using camera traps versus vehicle surveys, I performed an independent sample t-test. Since these two survey methods vary in their methods (e.g., stationary vs. mobile), I applied a Welch’s t-test because it does not assume equal variances between datasets (Student, 1908; Ruxton, 2006). The t-test was conducted in Microsoft Excel V. 365 (Microsoft Corporation, 2021) using its built-in “T.TEST” function with the parameter set for unequal variances (type = 3).

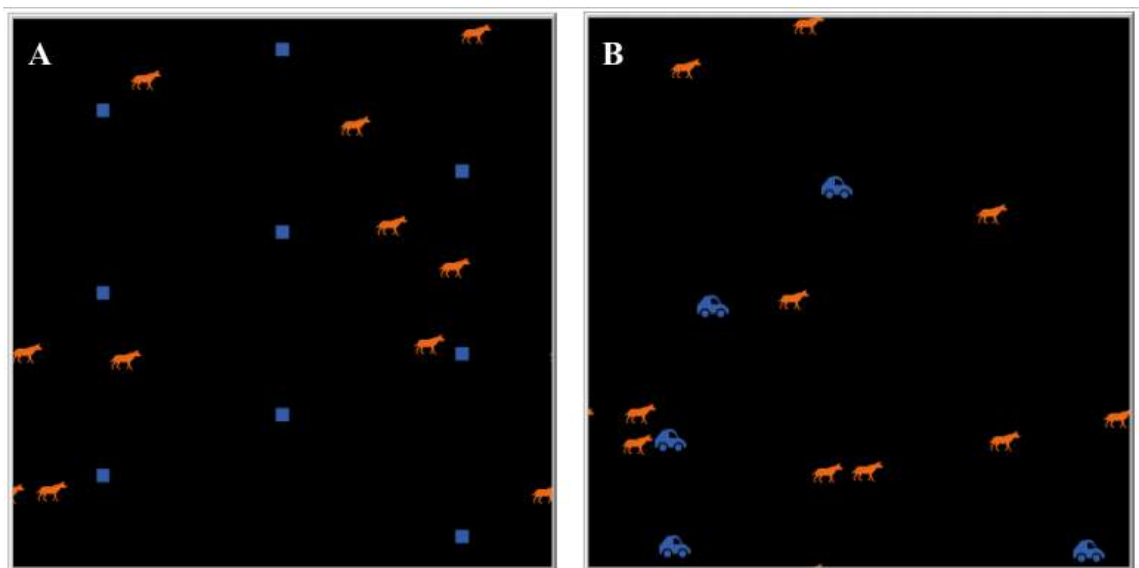


Figure 1. Initial NetLogo environments for the (A) camera trap survey and (B) vehicle survey. The orange ‘wolf’ icon represents the ten tigers moving at random. The blue squares (A) represent the nine stationary cameras evenly spaced in a grid. The blue vehicles (B) represent the five vehicles located and moving at random.

Results

Summary statistics from the NetLogo simulations demonstrated that camera traps required significantly less time than vehicle surveys to count the total number of tigers in an area. On average, camera traps required $42,926 \pm 19,609$ ticks (mean $\pm$ s.d.) with a minimum and maximum of 7,778–91,916 ticks. Alternatively, vehicle surveys required substantially more time to count the same number of tigers. On average, vehicle surveys required $259,104 \pm 87,614$ ticks (mean $\pm$ s.d.) with a minimum and maximum range of 109,250–475,094 ticks. The t-test confirmed that the difference in time between camera and vehicle surveys was highly significant: $t(32) = -13.2$, $p < 0.001$. Ultimately, these results indicate that, under the simulation conditions, camera traps are considerably more time-efficient than vehicle surveys for counting tigers.

Discussion

Scientifically accurate wildlife monitoring is important to track the status of any species as well as review the progress of conservation initiatives. The NetLogo results suggest that camera traps are significantly more time efficient in comprehensively counting wildlife counting as compared to vehicle surveys. Researchers and practitioners have often suggested that camera traps are the more efficient choice (O’Connell et al., 2011), and these results provide quantifiable evidence to support such claims. These results are critically informative because such a comparison is difficult to carry out in the field, particularly due to time and financial constraints that commonly limit conservation initiatives.

The time efficiency demonstrated in the NetLogo simulations may be due to the implementation of more camera traps ($n=9$) than vehicles ($n=5$) within the environment; in other words, cameras covered a wider area. However, the number of camera traps and vehicles were strategically selected to represent an intermediate number anticipated for use in each approach. If the initial environments were adjusted to hyperrealistic equipment quantities in conservation projects, then these results would likely be skewed even more in favor of camera traps; more camera traps (10-50 cameras/100km²) and fewer vehicles (2-10 per study area; O’Connell et al., 2011) would be initialized, thereby dramatically increasing the difference of an already statistically significant difference in time efficiency.

The time efficiency of camera traps may be attributed to their 24/7 operation that requires less human oversight and allows continuous monitoring, including overnight and in tough terrains where vehicles are limited (O’Connell et al., 2011). Although the setup and maintenance of camera traps requires upfront labor, this is typically offset overtime. Camera traps also offer enhanced detection efficiency via a more consistent and unbiased monitoring: motion-triggered photography. However, this method is not error-proof as misidentification can occur during data analysis (Karanth et al., 2006). The stationary placement of a camera also offers more reliable coverage compared to vehicle movement, which has the possibility of changing due to unpredictable field scenarios despite preestablished routes (O’Connell et al., 2011). The selection for camera traps can be attractive to certain researchers and conservationists because they are less intrusive to tiger behavior. However, some research contradicts this by suggesting that tigers

may avoid areas of recent human presence – including areas where camera traps were recently deployed (Karanth et al., 2006).

Alternatively, vehicle surveys require continuous human involvement, are typically restricted to daylight, and face physical challenges in rough terrain and complicated road networks (Karanth et al., 2006). These factors likely contribute to the larger variability and longer time requirements observed in the NetLogo vehicle survey simulations. Notably, vehicle surveys incur ongoing costs related to fuel and vehicle maintenance. They are also more susceptible to sampling errors (e.g., missing or double-counting tigers; Kerley et al., 2002). Altogether, fluctuating costs and higher error rates set forth substantial risks that budget-limited conservation initiatives may need to avoid. Yet, despite the apparent advantages of camera traps, vehicle surveys remain crucial in certain contexts, such as real-time visual assessments or when immediate intervention is required, such as in veterinary care for injured tigers. Further, vehicle surveys may be more advantageous when attempting to gather data on behavior, territory mapping, and ecological context that camera traps cannot provide. As technological advances continue to enhance both camera trap hardware and data analysis software, we anticipate that the gap in efficiency between the two methods will widen, further reinforcing the advantages of camera traps in tiger monitoring.

Future research should continue to discern between the use of camera traps versus vehicle surveys: *In what context?* And *How to optimize the deployment of each survey type?* Even before *in situ* research takes place, future agent-based modeling could account for habitat-specific field conditions. For example, vegetation type and tiger behavior could be incorporated to more accurately simulate tiger movement and counting, ultimately yielding more realistic time estimates which could assist practitioners in project schedules and budgets. Research on camera trap surveys should focus on maximizing their coverage in a cost-effective manner (e.g., perhaps by finding the optimal number of camera traps per area, beyond which more are not needed). This may include comparisons of difficult camera models or arrangements in the field (e.g., height, angle, etc.). Notably, there should be continued research on the utility of hybrid methods that combine the strengths of both vehicle surveys and camera traps: *How can data from each type of survey be synergistic?* Additionally, studies investigating the social implications of these methods, such as community acceptance and engagement with technology, could further refine conservation strategies.

Overall, the results strongly favor the use of camera traps over vehicle surveys for time-efficient tiger monitoring. While both methods have their place in conservation, camera traps are promising for their ability to provide accurate, consistent data with minimal human effort and animal disturbance (O’Connell et al., 2011). Nevertheless, vehicles should be considered when real-time observations are advantageous (e.g., veterinary visual assessments). Additionally, while some vehicle surveys have led to tiger avoidance of an area, they can also lead to tiger habituation which may be advantageous for observational research and tourism (Kerley et al., 2002). These results have important implications for wildlife conservation because they demonstrate that camera traps can provide timely and reliable wildlife count data. Now, the

Indian government is starting to incorporate AI to process collected images. This will ensure that camera traps are quick and effective not only in the field, but also in data analysis. The findings from this study are critical for those involved in India's tiger conservation: For PT, who may have the power to purchase and distribute camera traps widely; For NGOs who may work in logistically challenging unprotected tiger habitat; For community-based initiatives who may be limited in labor or funds. The strength of all sectors collaborating, and optimizing their methodology, allows for the future of tiger monitoring to continue in a consistent, accurate manner thereby promoting a strong understanding of tiger population trends and conservation trajectories.

Acknowledgements

Thank you Dr. Mary Burak for the mentorship throughout all stages of this project, from study design to paper writing.

Works Cited

- Arjunan, M., C. Holmes, J. Puyravaud, and P. Davidar. "Do Developmental Initiatives Influence Local Attitudes toward Conservation? A Case Study from the Kalakad-Mundanthurai Tiger Reserve, India." *Journal of Environmental Management*, vol. 79, no. 2, Apr. 2006, pp. 188–197, doi:10.1016/j.jenvman.2005.06.007.
- Chowdhury, A. N., et al. "Stigma of Tiger Attack: Study of Tiger-Widows from Sundarban Delta, India." *Indian Journal of Psychiatry*, vol. 58, no. 1, 2016, pp. 12–19.
- Diepstraten, J., et al. "Assessing Project Proposals Based on National and Global Tiger Action Plans: Lessons from the Integrated Tiger Habitat Conservation Programme (ITHCP)." *Land*, vol. 11, no. 12, 2022, 2326.
- Goodrich, J., et al. "Panthera tigris." *The IUCN Red List of Threatened Species*, 2022, e.T15955A214862019, doi:10.2305/IUCN.UK.2022-1.RLTS.T15955A214862019.
- Gopalaswamy, A. M., et al. "How 'Science' Can Facilitate the Politicization of Charismatic Megafauna Counts." *Proceedings of the National Academy of Sciences*, vol. 119, no. 20, 2022, e2203244119, doi:10.1073/pnas.2203244119.
- Jhala, Y., et al. "Recovery of Tigers in India: Critical Introspection and Potential Lessons." *People and Nature*, vol. 3, no. 2, 2021, pp. 281–293.
- Karanth, K. U., et al. "Assessing Tiger Population Dynamics Using Photographic Capture-Recapture Sampling." *Ecology*, vol. 87, no. 11, 2006, pp. 2925–2937.
- Kerley, L. L., et al. "Effects of Roads and Human Disturbance on Amur Tigers." *Conservation Biology*, vol. 16, no. 1, 2002, pp. 97–108.
- Microsoft Excel. Version 365, Microsoft Corporation, 2021.
- Ministry of Environment, Forest and Climate Change. "Achievements of National Tiger Conservation Authority (NTCA) during the Year 2023." Press Information Bureau, 29 Dec. 2023, pib.gov.in/PressReleasePage.aspx?PRID=1991620.
- O'Connell, A. F., J. D. Nichols, and K. U. Karanth. *Camera Traps in Animal Ecology: Methods and Analyses*. Springer, 2011.
- Railsback, S. F., and V. Grimm. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton University Press, 2019.
- Ruxton, G. D. "The Unequal Variance T-Test Is an Underused Alternative to Student's T-Test and the Mann–Whitney U Test." *Behavioral Ecology*, vol. 17, no. 4, 2006, pp. 688–690.
- Sethi, S. "Insights into Illegal Wildlife Hunting by Forest Guards of Selected Tiger Reserves in Central India." *European Journal of Wildlife Research*, vol. 68, no. 4, 2022, <https://doi.org/10.1007/s10344-021-01553-8>.
- Sinha, S., et al. "Tigers Losing Grounds: Impact of Anthropogenic Occupancy on Tiger Habitat Suitability Using Integrated Geospatial-Fuzzy Techniques." *The Ecoscan*, vol. 1, 2012, pp. 259–263.
- Student. "The Probable Error of a Mean." *Biometrika*, vol. 6, no. 1, 1908, pp. 1–25.

- Wilensky, U. NetLogo [Computer software]. Center for Connected Learning and Computer-Based Modeling, Northwestern University, 1999, <http://ccl.northwestern.edu/netlogo/>.
- Wilensky, U., and W. Rand. An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. MIT Press, 2015.
- Yumnam, B., et al. "Prioritizing Tiger Conservation Through Landscape Genetics and Habitat Linkages." PLOS ONE, vol. 9, no. 11, 2014, e111207.

A Regression Study of EV in Developed Countries By Yijie Fu

Abstract

This paper examines electric vehicle (EV) adoption in developed countries using multi-variable linear regression. Analyzing data from 12 nations spanning 2016–2023, the study investigates correlations between EV adoption ratios and four key factors: median age, average disposable income, the share of renewable energy in primary energy supply, and per capita primary energy supply. Unlike prior research that emphasizes financial incentives (Sierzchula, 2013) and charging infrastructure, this study integrates socio-economic and demographic variables, providing a broader analytical scope.

Findings reveal strong positive correlations between EV adoption and renewable energy utilization and average disposable income, while median age and per capita primary energy supply exhibit weaker associations. The study employs rigorous linear regression techniques, detailing one- and multiple-parameter models. Examples demonstrate the derivation of regression coefficients, loss function minimization, and the role of the design matrix and Hessian matrix in achieving optimal solutions. The significance of the r^2 coefficient in evaluating model performance is underscored.

The results underscore the pivotal roles of economic prosperity and the share of renewable energy in driving EV adoption, offering actionable insights for policymakers. Recommendations include subsidizing renewable energy generating firms, and promoting public environmental awareness. This research also highlights areas for future exploration, such as charging infrastructure and cross-country policy analyses, to advance global sustainable transportation initiatives.

Keywords

Electric Vehicles, EV Adoption, Linear Regression, Renewable Energy, Disposable Income, Socio-Economic Factors, Sustainable Transportation, Policy Analysis, Data Analytics, Developed Countries.

1 Introduction: Tracing the Origins and Purpose of Linear Regression in EV Adoption Research

The history of linear regression dates back to the early 19th century and is tied to the development of the least squares method. It began with Adrien-Marie Legendre (1805), who introduced the method of least squares to estimate parameters in fitting a curve to observed data. Carl Friedrich Gauss, around the same time, expanded on Legendre's work and formalized the method, applying it to astronomy for predicting planetary movements (Seber & Lee, 2012; Stigler, 1986).

Gauss is credited with the introduction of the normal distribution, which became essential in understanding the errors in predictions and how the least squares method minimize them. By the mid-19th century, Sir Francis Galton further advanced the regression analysis by introducing

the term “regression” in his work on heredity. He observed that offspring’s traits tended to regress toward the average, leading to the development of the concept of regression to the mean. In the early 20th century, Karl Pearson expanded on Galton’s work by formalizing the relationship between variables, leading to the development of correlation coefficients. Further advances came with the work of Ronald A. Fisher, who introduced the concept of statistical significance testing and helped establish the foundations of modern linear regression analysis as part of a broader statistical framework (Seber & Lee, 2012; Stigler, 1986).

In summary, the history of linear regression spans from early applications in astronomy to its formalization in genetics and statistics, making it a cornerstone of statistical analysis today. This study employs linear regression to examine the factors influencing the adoption of electric vehicles (EVs) in developed countries. This area of research is increasingly relevant in light of global efforts to reduce carbon emissions and advance the transition to sustainable energy systems. EV adoption is a key measure of progress in these initiatives, yet understanding the underlying drivers remains a complex challenge. The analysis focuses on four independent variables—median age, average disposable income, the percentage of renewable energy in total primary energy supply, and primary energy supply per capita. These factors were chosen because they encapsulate significant demographic, economic, and energy-related dimensions that are likely to influence EV adoption. Investigating these relationships provides a basis for identifying patterns that can guide policy decisions aimed at accelerating EV adoption and fostering long-term environmental sustainability.

The rest of the paper is organised as follows. In Section 2, we explain how to perform linear regression with one independent variable and show that linear regression minimizes the sum of squared errors. Then in Section 3, we generalize to the case of multiple independent variables. In Section 4 we introduce an important metric for studying the strength of correlation, the r^2 -coefficient. Finally, in Section 5, we use the developed methods to investigate the correlation between Median age average disposable income, the percentage of renewable energy in total primary energy supply, primary energy supply per capita as the independent variables, and the adoption of electric vehicles (EVs) as the dependent variable. The calculation indicates that average disposable income (in USD) and the percentage of renewable energy in total primary energy supply were highly correlated within the EV ratio, which means a shift in them is likely to cause a big change in Ev ratio.

2 One-parameter linear regression: Unraveling the Mechanics and Validation

This section focuses on one-parameter linear regression. It details fitting a line ($Y = \beta_0 + \beta_1 X$) derived through partial differentiation. The optimality of the obtained coefficients is verified using the Hessian matrix and the second derivative test. An example with given data points [(3, 5), (7, 11), (8, 16), (10, 20)] showcases the practical application, computing sums and substituting into formulas to obtain the regression line $Y = -2.05 + 2.15X$.

2.1 Fitting the Line to the Data Points: Estimating Regression Coefficients

In simple linear regression, we aim to model the relationship between a dependent variable Y and an independent variable X using a straight line. The model can be expressed as:

$$Y = \beta_0 + \beta_1 X$$

where:

- Y is the dependent variable,
- X is the independent variable,
- β_0 is the intercept (constant term),
- β_1 is the slope (coefficient of X),

The parameters β_0 and β_1 are estimated by minimizing the sum of squared errors (loss function), which is given by:

$$L(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2$$

where n is the number of observations, y_i is the observed value of Y for the i-th observation, and x_i is the corresponding value of X.

The values of β_0 and β_1 that minimize the loss function can be found using partial differentiation. After partial differentiation in terms of β_0 and β_1 respectively, and letting the results both equal to zero, we obtain a system of two simultaneous equations which implies there is a unique solution (β_0, β_1) . Let us do the general computation to obtain the formulas for β_0 and β_1 .

First, we take the partial derivative of $L(\beta_0, \beta_1)$ with respect to β_0 :

$$\frac{\partial L(\beta_0, \beta_1)}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)$$

Setting this derivative equal to zero to find the minimum:

$$\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

Simplifying:

$$\sum_{i=1}^n y_i = n\beta_0 + \beta_1 \sum_{i=1}^n x_i$$

This gives the first equation (2.1):

$$\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n}$$

Next, we take the partial derivative of $L(\beta_0, \beta_1)$ with respect to β_1 :

$$\frac{\partial L(\beta_0, \beta_1)}{\partial \beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i$$

Setting this derivative equal to zero to find the minimum:

$$\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

Substitute β_0 from equation (2.1) into

$$\sum_{i=1}^n y_i x_i = \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2$$

We get

$$\sum_{i=1}^n y_i x_i = \frac{\sum y_i - \beta_1 \sum x_i}{n} \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2$$

This simplifies to

$$\beta_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \left(\sum x_i \right)^2}$$

which is the formula for β_1 .

Substitute the value of β_1 back into equation (2.1) to find β_0 :

$$\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n}$$

Thus, we've derived the formulas for β_0 and β_1 using partial differentiation.

$$\beta_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \left(\sum x_i \right)^2}$$

$$\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n}$$

These values minimize the sum of squared errors, providing the least squares estimates for the linear regression coefficients.

2.2 Verifying the Optimality: Linear Regression Produces a Minimum

We still have to check that the obtained point (β_0, β_1) is a minimum of the loss function.

Theorem 2.1

If $f(x, y)$ is a twice-differentiable function of two variables and (x_0, y_0) is a solution of the system

$$\frac{\partial f}{\partial x} = 0 \text{ and } \frac{\partial f}{\partial y} = 0$$

then the point (x_0, y_0) is a minimum if the determinant of the Hessian matrix

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

is positive and $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$.

Proof: See [Calculus: Early Transcendental by James Stewart](#) for the proof.

Proposition 2.2. The point (β_0, β_1) obtained by linear regression is a minimum.

Proof: We follow the proof in [Theorem 2.1](#) (Simon & Blume, 2007; Apostol, 1969) first-order partial derivatives of $S(\beta_0, \beta_1)$ are:

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)$$

$$\frac{\partial S}{\partial \beta_1} = -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i)$$

Now, we compute the second-order partial derivatives:

$$\frac{\partial^2 S}{\partial \beta_0^2} = 2n$$

$$\frac{\partial^2 S}{\partial \beta_0 \partial \beta_1} = 2 \sum_{i=1}^n x_i$$

$$\frac{\partial^2 S}{\partial \beta_1^2} = 2 \sum_{i=1}^n x_i^2$$

Thus, the Hessian matrix H is:

$$H = \begin{pmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{pmatrix}$$

Let us check the second derivative test conditions:

The determinant of the Hessian is:

$$\det(H) = 2n \cdot 2 \sum_{i=1}^n x_i^2 - \left(2 \sum_{i=1}^n x_i \right)^2$$

Simplifying:

$$\det(H) = 4 \left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right)$$

Note that the term $n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2$ is the variance of the x_i 's multiplied by n . Thus, it is positive, and the determinant is positive.

2.3 Implementing the method: Example of One-Parameter Linear Regression

Suppose that we are given the following data points:

(3, 5), (7, 11), (8, 16), (10, 20)

We want to fit a linear model of the form:

$$Y = \beta_0 + \beta_1 X$$

Calculate the necessary sums For each data point (x_i, y_i) , calculate the following quantities:

$$\sum x_i = 3 + 7 + 8 + 10 = 28$$

$$\sum y_i = 5 + 11 + 16 + 20 = 52$$

$$\sum x_i y_i = 3 \times 5 + 7 \times 11 + 8 \times 16 + 10 \times 20 = 15 + 77 + 128 + 200 = 420$$

$$\sum x_i^2 = 3^2 + 7^2 + 8^2 + 10^2 = 9 + 49 + 64 + 100 = 222$$

To solve for β_0 and β_1 we substitute into the formulas in the proof of [Proposition 2.2](#) for the least squares estimates of β_0 and β_1 :

$$\beta_1 = \frac{4 \times 420 - 28 \times 52}{4 \times 222 - 28^2} = \frac{1680 - 1456}{888 - 784} = \frac{224}{104} \approx 2.15$$

$$\beta_0 = \frac{52 - 2.15 \times 28}{4} = \frac{52 - 60.2}{4} = \frac{-8.2}{4} = -2.05$$

Thus, the estimated regression line is:

$$Y = -2.05 + 2.15 X$$

3 Extending the Paradigm: Multiple-parameter linear regression

In this section, the multiple linear regression model ($y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$) is explored. The loss function is defined, and partial differentiation wrt each coefficient yields a system of equations. The design matrix X leads to the derivation of normal equations ($X^T X \beta = X^T Y$), with the solution $\beta = (X^T X)^{-1} X^T Y$ when $X^T X$ is invertible. The Hessian matrix $H = 2X^T X$ is analyzed; its positive definiteness, linked to $X^T X$'s column independence, ensures the loss function's minimum. An example of two-parameter regression on a dataset demonstrates the process, computing X , Y , solving equations to get coefficients, and forming the model $\hat{y} = 1.0 + 0.5x_1 + 1.5x_2$

The general form of a multiple linear regression model (3.1) is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

where:

- y is the dependent variable.
- $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ are the coefficients (parameters) of the model.

- $x_1, x_2, \dots, x_n$ are the independent variables.

Suppose that we have observations $(x_{i1}, x_{i2}, \dots, x_{in}, i = 1, \dots, m)$ of the independent variables. We let:

$$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in}$$

Similarly to the case of one independent variable, we define the loss function (3.2) as:

$$L(\beta_0, \dots, \beta_n) = \sum_{i=1}^m (y_i - \hat{y}_i)^2 = \sum_{i=1}^m (y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in}))^2$$

3.1 Applying Partial Differentiation: Deriving the System of Equations:

We want to minimize the loss function with respect to the coefficients $(\beta_0, \beta_1, \beta_2, \dots, \beta_n)$.

To do this, we take the partial derivative of $L(\beta_0, \dots, \beta_n)$ with respect to each of these coefficients and set the derivatives equal to zero.

Partial Derivative with Respect to β_0 :

$$\frac{\partial L}{\partial \beta_0} = -2 \sum_{i=1}^m \left(y_i - \beta_0 - \sum_{j=1}^n \beta_j x_{ij} \right) = 0$$

Partial Derivative with Respect to β_1 :

$$\frac{\partial L}{\partial \beta_1} = -2 \sum_{i=1}^m \left(y_i - \beta_0 - \sum_{j=1}^n \beta_j x_{ij} \right) x_{i1} = 0$$

Partial Derivative with Respect to β_2 :

$$\frac{\partial L}{\partial \beta_2} = -2 \sum_{i=1}^m \left(y_i - \beta_0 - \sum_{j=1}^n \beta_j x_{ij} \right) x_{i2} = 0$$

This process continues for each β_j until β_n .

After differentiating the loss function with respect to each β_j , we set each equation to zero to find the optimal values of the coefficients that minimize the loss function.

This yields a system of simultaneous equations:

$$\begin{aligned} \sum_{i=1}^m (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_n x_{in}) &= 0 \\ \sum_{i=1}^m (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_n x_{in}) x_{i1} &= 0 \\ \sum_{i=1}^m (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_n x_{in}) x_{i2} &= 0 \\ &\vdots \end{aligned}$$

$$\sum_{i=1}^m (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_n x_{in}) x_{in} = 0$$

3.2 Solving the System of Equations: Employing Matrix Algebra

Definition 3.1. The design matrix X is a matrix that contains the independent variables (input data) for the regression model. Each row represents an observation, and each column represents an independent variable (including an intercept term). Explicitly, we have

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1n} & 1 & x_{21} & x_{22} & \dots & x_{2n} & \vdots & \vdots & \vdots & \ddots & \vdots & 1 & x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

Definition 3.2. For a given matrix X , its transpose, denoted as X^T is formed by turning its rows into columns and its columns into rows.

The system of equations above can be written in matrix form as the normal equations as we explain below:

$$X^T X \beta = X^T Y$$

where:

- X is the design matrix of the independent variables $x_{i1}, x_{i2}, \dots, x_{in}$,
- Y is the vector of observed values $y_1, y_2, \dots, y_m$,
- β is the vector of unknown coefficients $\beta_0, \beta_1, \dots, \beta_n$.

(Seber & Lee, 2012; Montgomery et. al., 2015)

To find the best-fit linear model, we minimize the function $L(\beta_0, \dots, \beta_n)$ from (3.2). In the matrix form, it can be expressed as:

$$L(\beta) = (Y - X\beta)^T (Y - X\beta)$$

To minimize $L(\beta)$, take the derivative of the loss function with respect to β and set it equal to zero:

$$\frac{\partial}{\partial \beta} L(\beta) = -2X^T(Y - X\beta) = 0$$

This leads to the normal equation:

$$X^T X \beta = X^T Y$$

If the matrix $X^T X$ is invertible (i.e., X has full rank), the solution for the regression coefficients β is:

$$\beta = (X^T X)^{-1} X^T Y$$

This gives the coefficients $(\beta_0, \beta_1, \dots, \beta_n)$ that minimize the sum of squared errors and provide the best fit for the linear regression model.

So the solution to this system is:

$$\beta = (X^T X)^{-1} X^T Y$$

This gives the values of $(\beta_0, \beta_1, \dots, \beta_n)$ that minimize the loss function.

To show that the linear regression produces a minimum, we need to first have a generalization of Theorem 2.1 to the case of $n + 1$ variables.

Suppose that we have a function $f(x_1, \dots, x_{n+1})$ that is twice differentiable. We need to introduce some linear algebra concepts first.

Definition 3.3. A matrix A is positive definite if and only if for any non-zero vector z , we have $z^T A z > 0$

Theorem 3.4. If the Hessian matrix X positive definite, then the values of $\beta_0, \beta_1, \dots, \beta_n$ is the value that minimizes the loss function

Then we define the Hessian matrix H as the matrix of second-order partial derivatives of the function $f(\beta)$ with respect to the regression coefficients $\beta_0, \beta_1, \dots, \beta_n$:

$$H = \frac{\partial^2 f}{\partial \beta_i \partial \beta_j}$$

3.3 Analyzing Convexity: proving positive definiteness of Hessian Matrix in Multiple Linear Regression

For linear regression, the Hessian matrix is given by:

$$H = 2 X^T X$$

$$H = 2 \begin{pmatrix} m \sum_{i=1}^m x_{i1} & m \sum_{i=1}^m x_{i2} & \dots & m \sum_{i=1}^m x_{in} & m \sum_{i=1}^m x_{i1}^2 & m \sum_{i=1}^m x_{i1} x_{i2} & \dots & m \sum_{i=1}^m x_{i1} x_{in} & m \sum_{i=1}^m x_{i2}^2 & m \sum_{i=1}^m x_{i2} x_{i1} & \dots & m \sum_{i=1}^m x_{i2}^2 & \dots & m \sum_{i=1}^m x_{i2} x_{in} \end{pmatrix}$$

where X is the design matrix of the independent variables.

(Goldfarb, 1970; Montgomery et. al., 2015)

As discussed in the previous sections, the normal equation for multiple linear regression is given by:

$$X^T X \beta = X^T Y$$

where:

- X is the design matrix of the independent variables,
- Y is the vector of observed values,
- β is the vector of coefficients $\beta_0, \beta_1, \dots, \beta_n$.

To solve for β , we require $X^T X$ to be invertible, and this requires the matrix $X^T X$ to be positive definite. We now prove that $X^T X$ is positive definite when the columns of X are linearly independent.

By Definition 3.3, we need to show that:

$$z^T (X^T X) z > 0 \text{ for any non-zero vector } z.$$

Let $z^T (X^T X) z = (Xz)^T (Xz)$ be a non-zero vector. The expression $z^T (X^T X) z$ can be rewritten as:

$$z^T(X^T X)z = (Xz)^T(Xz)$$

Here, Xz is the result of multiplying matrix X by vector z . The expression $(Xz)^T(Xz)$ is the dot product of Xz with itself, which yields the squared norm of Xz :

$$(Xz)^T(Xz) = |Xz|^2$$

Because the columns of X are linearly independent, the matrix X has full rank. Therefore, for any non-zero vector z , the vector Xz will also be non-zero. Consequently, the norm $||Xz||^2$ is positive for any non-zero z :

$$||Xz||^2 > 0.$$

Thus, we have:

$$z^T(X^T X)z = |Xz|^2 > 0.$$

Since $z^T(X^T X)z > 0$ for any non-zero z , we conclude that $X^T X$ is positive definite, meaning all its eigenvalues are positive. This positive definiteness implies that the Hessian matrix $H = 2X^T X$, as introduced earlier, is also positive definite. Therefore, the solution to the normal equation corresponds to a minimum of the loss function in multiple linear regression.

Since the Hessian matrix is positive definite, loss function is convex. This proves that the values obtained by solving the linear regression problem correspond to the minimum of the variance.

3.4 Demonstrating Application: Example for Multiple-parameter Linear Regression

In this example, we perform a two-parameter linear regression using the dataset:

x_1	x_2	y
1	2	3
2	3	6
3	4	7
4	5	8
5	6	11
6	7	13

Table1: Dataset for example of Multiple-parameter Linear Regression

We aim to fit the model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$$

The loss function is given by:

$$L(\beta_0, \beta_1, \beta_2) = \sum_{i=1}^6 (y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))^2$$

$$X = (1 \ 1 \ 2 \ 1 \ 2 \ 3 \ 1 \ 3 \ 4 \ 1 \ 4 \ 5 \ 1 \ 5 \ 6 \ 1 \ 6 \ 7) \quad Y = (3 \ 6 \ 7 \ 8 \ 11 \ 13)$$

The normal equations are:

$$X^T X \beta = X^T Y$$

After computing $X^T X$ and $X^T Y$, we have:

$$(6 \ 21 \ 27 \ 21 \ 91 \ 112 \ 27 \ 112 \ 139) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = (48 \ 194 \ 241)$$

Solving the system of equations, we get the coefficients:

$$\beta_0 = 1.0, \quad \beta_1 = 0.5, \quad \beta_2 = 1.5$$

The final regression model is:

$$\hat{y} = 1.0 + 0.5x_1 + 1.5x_2$$

4 Assessing Model Fit: The R^2 Coefficient

In the context of multiple linear regression, the R^2 coefficient, also known as the coefficient of determination, plays a crucial role in evaluating the goodness of fit of the regression model. It provides a measure of how well the independent variables explain the variability in the dependent variable.

Definition of R^2

The R^2 coefficient is defined as:

$$R^2 = 1 - \frac{RSS}{TSS}$$

Where loss function is the residual sum of squares:

$$loss \ function = \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

This measures the total squared deviation of the predicted values $\hat{y}_i$ from the actual observed values y_i , representing the unexplained variability by the model.

TSS is the total sum of squares:

$$TSS = \sum_{i=1}^m (y_i - \bar{y})^2$$

where $\bar{y}$ is the mean of the observed values y_i . This measures the total variability in the observed data.

Thus, R^2 can be interpreted as the proportion of the total variability in the dependent variable Y that is explained by the independent variables in the model.

Role of R^2 in Regression

The R^2 value ranges between 0 and 1:

- $R^2 = 1$ indicates that the regression model perfectly explains the variability in the dependent variable, meaning that all observed data points lie exactly on the predicted regression line.
- $R^2 = 0$ means that the model does not explain any of the variability in the dependent

variable, implying that the independent variables have no linear relationship with the dependent variable.

In practice, the R^2 coefficient provides an indication of how well the regression model fits the data:

$$0 \leq R^2 \leq 1$$

The closer R^2 is to 1, the better the model explains the variation in the dependent variable.

5 Unveiling EV Adoption Dynamics: Practical study of linear regression

In this section, the paper applied the principles of multiple-parameter linear regression to examine factors influencing the adoption of electric vehicles (EVs) in developed countries. The study seeks to identify key drivers of EV adoption by exploring correlations between the EV ratio and four variables: average disposable income, the percentage of renewable energy in the total primary energy supply, primary energy supply per capita, and median age. Below is the calculated correlation matrix.

	EV Adoption	Disposable Income	Renewable Energy	Energy Supply	Median Age
EV Adoption	1.0000	0.4455	0.7591	0.0935	-0.1210
Disposable Income	0.4455	1.0000	0.6723	0.3851	0.0322
Renewable Energy	0.7591	0.6723	1.0000	0.5876	-0.1805
Energy Supply per Capita	0.0935	0.3851	0.5876	1.0000	0.0509
Median Age	-0.1210	0.0322	-0.1805	0.0509	1.0000

Table 1: The correlation matrix of EV adoption and other factors

5.1 Economic Prosperity and EV Adoption

Our research has unearthed a significant link between average disposable income (Pearson coefficient 0.4455) and EV adoption. Sierzechula et al. (2013) previously probed the role of financial incentives, but we've delved deeper into the income-EV relationship. In developed countries, this discovery signals policymakers to devise customized economic strategies. Tailored subsidies, aligned with income levels, can supercharge EV adoption. By zeroing in on those for whom EVs are nearly affordable, we can trigger a surge in purchases. Middle-income consumers, often daunted by EVs' heftier price tags, could be swayed by astutely designed incentives. This approach capitalizes on our finding and forges a more inclusive EV market, a dimension not fully explored before.

5.2 Renewable Energy Infrastructure and EV Adoption

The strong correlation (Pearson coefficient 0.7591) between renewable energy percentage and EV uptake is revolutionary. Wang et al.'s (2016) global analysis offered an overview, yet our

study spotlights actionable policies. In developed countries, mandating renewable energy for EV charging is imperative. We suggest policies ensuring charging stations source a portion of power from renewables. As EVs multiply, so does the appetite for clean energy, creating a virtuous loop. Public-private partnerships can fuel investment in both EV charging grids and renewable generation. Our discovery paves the way for a seamless, sustainable energy-vehicle ecosystem, a unique emphasis of our research.

5.3 Limited Influence of Demographic and Energy Supply Variables

The weak correlation between EV adoption and median age (-0.1210) debunks age-driven adoption theories. In developed nations, our finding dismantles the myth of generational determinism. Policy efforts should pivot to universal environmental messaging. For primary energy supply per capita (0.0935), the weak tie shows infrastructure outweighs supply volume. Our research champions strategic urban planning. Charging points must be woven into the urban tapestry, from business districts to neighborhoods. This data-driven approach eclipses earlier, broader strokes.

5.4 Implications for Policy and Future Research

Our findings are a guiding light for policymakers. Beyond economic and energy measures, regulatory frameworks need a revamp. Emissions standards can be recalibrated to favor EVs, spurring market demand. Future research, inspired but distinct from predecessors, should plunge into cultural terrain. How do cultural views on environmental care interact with EV incentives? Cross-country studies, dissecting cultural idiosyncrasies, can unearth new adoption boosters. By anchoring on our insights, developed countries can pioneer policies, heralding a new era of sustainable EV mobility.

Here is the dataset that I collected and used: the data of the EV regression study

Works Cited

- Goldfarb, D. A family of variable - metric methods derived by variational means. *Mathematics of Computation*, 1970, 24: 23 - 26.
- Seber, G. A. F., Lee, A. J. *Linear Regression Analysis*. Wiley, 2012.
- Stigler, S. M. *The History of Statistics: The Measurement of Uncertainty before 1900*. Harvard University Press, 1986.
- Simon, C. P., Blume, L. *Mathematics for Economists*. W.W. Norton & Company, 2007.
- Apostol, T. M. *Calculus, Volume 2: Multi - Variable Calculus and Linear Algebra*. Wiley, 1969.
- Montgomery, D. C., Peck, E. A., Vining, G. G. *Introduction to Linear Regression Analysis*. Wiley, New York, NY, 2015.
- Sierzchula, W., Bakker, S., Maat, K., van Wee, B. The influence of financial incentives and charging infrastructure on electric vehicle adoption. *Transportation Research Part D: Transport and Environment*, 2013, 25: 39 - 49.

Air Pollution and Air Quality Index(AQI) Trends and Predictions Using AI Model

By Rishan Pawar

Introduction

The motivation behind this study is to understand the trends of and forecast air quality of Colorado. There are many purposes that make this research important, specifically to help the citizens that live in the polluted state. Colorado is known for its thinner air because it is thousands of feet above sea level. In addition, its air quality levels can deplete when cars emit harmful gasses, like ozone precursors ([REFERENCE](#)). The individuals are doing the same habits as those in other states; however, because the air quality is already poor to begin with, such Colorado can help fix health implications and negative environmental impacts. In addition it can cause public awareness and fix regulations and policies.

The health implications can cause lethal diseases and actually lead to a variety of health problems in the future as well. In Colorado it is more likely for the average citizen to receive lung cancer, cardiovascular diseases, or asthma. In fact, more than half of the deaths from the 1990's to 2020 were because of cardiovascular and lung diseases([REFERENCE](#)). In addition, there is an increased chance of premature death. This can be because, again, citizens have trouble breathing or those specific individual's have lung disease that have not been detected. Negative environment factors can damage crops, meaning a lack of food, harm ecosystems, and can be a major contributor to global warming. Even the geography of Colorado is unique and very vulnerable to wildfires. The dry climate, mountainous terrain, and droughts can make perfect conditions for wildfires.

There are numerous precautions that people can take in Colorado, yet some of them are not completely aware of the situation. Therefore, researchers or experts in this study want to spread public awareness. Just to list a few, researchers might, in order to benefit the citizens of Colorado, try to protect health, make informed decisions, advocate for clean air, or even take action and protect people and the general environment. In order to protect their health, they can limit habits, such as outdoor activities and stay indoors during the days that say that the AQI will be high, but if they have to go outside for a work purpose or school, they can wear masks. In addition people can start spreading public awareness by advocating for it. These methods can protect both people and the environment.

Our research can create public awareness for the terrible air quality. This is to help, as mentioned before, the people in Colorado. Policies that help improve the air quality can make a healthier population, stronger economy, enhanced quality of life, environmental protection, and all of these can cause reduced healthcare costs. The policies can prevent diseases and the prevention of disease can cause a healthier lifestyle. In fact, the healthier population can boost the local economy by attracting business and tourists. When there are policies, there can be an enhanced quality of life, meaning that residents of Colorado can participate in outdoor activities like other US citizens. By doing this research, we are trying to better the life of every Colorado resident, make the opportunities more broad.

The objective of this study is to understand pattern, trends, and seasonality of air quality, specifically ozone pollution in Denver, Colorado. Additionally, we aim to develop a time series machine learning model to predict ozone pollution into the near future in Denver, Colorado.

Dataset

In this study we used air pollution data in the United States from 2000 to 2021([REFERENCE](#)). Our dataset includes the daily sampled ozone (O3) mean, O3 air quality index (AQI), and other pollution metrics from different states and cities. We have filtered our dataset to Denver, Colorado. Due to the missing values in the early 2000's, we used the historical ozone data from 2012 to 2021, as shown in Figure 1.

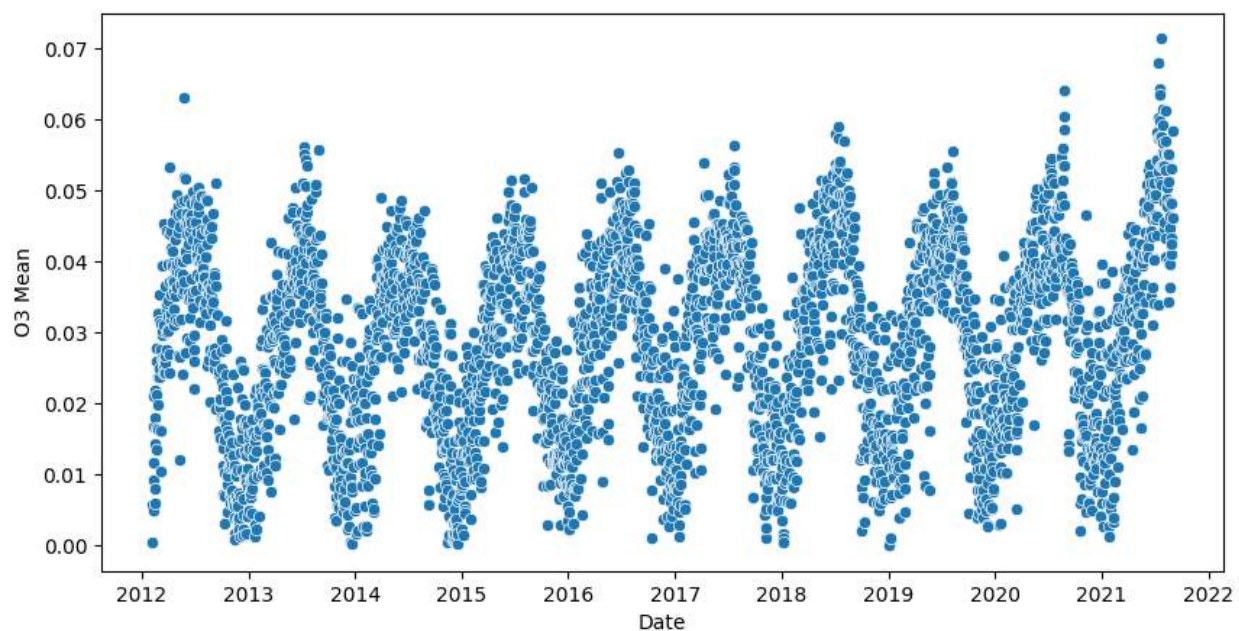


Figure 1. Daily sampled historical ozone mean data

Methods and Models

Stationarity Test

First, we tested our dataset for stationarity. Stationarity is a crucial property in time series analysis. A stationary time series has statistical properties (mean, variance, autocorrelation) that remain constant over time. The Augmented Dickey-Fuller (ADF) test is used to check whether a time series is stationary. The test has two hypotheses: the null hypothesis (h_0), which assumes that the data is not stationary, and the alternative hypothesis (h_1), which assumes that the data is stationary. The test produces a p-value, which tells us how likely the data is to follow the null hypothesis. The p-value is very small (0.003361) and is well below the typical significance levels. This indicates that we can reject the null hypothesis (which assumes non-stationarity) with a high degree of confidence. The autocorrelation function (ACF) statistics of our dataset is significantly less than the critical values at all levels (1%, 5%, and 10%), which further supports

that the data is stationary. This is a positive sign for many time series analysis techniques, as stationarity is often a prerequisite.

Autocorrelation and Partial Autocorrelation Functions

Autocorrelation functions (ACF) and partial autocorrelation functions (PACF) are both tools used in time series data to measure the correlation between a variable at one point in time and its values at earlier times. ACF measures the correlation between a variable at time t and its lagged values at previous time points ($t-1$, $t-2$...). A positive ACF means that if the variable is high at a certain time, it will likely be high at the next time point too. A negative ACF means that if the variable is high now, it's likely to be low after a certain period. A decaying ACF means that the relationship weakens over time; the influence of past values is stronger at earlier lags and fades as the time gap increases.

PACF is similar to ACF, but it measures the correlation while controlling the influence of values in between. This means it isolates the relationship between today's value and a specific past value, without being affected by the values between them. A significant PACF at a particular lag suggests a direct relationship between the variable at t and a certain value from one time ago. Often, the PACF cuts off after a certain lag, indicating that the direct relationship between the variable and its lagged values beyond that lag is negligible.

In this study, ACF and PACF as shown in Figure 2 were used to understand relationships between ozone levels and their lagged values. As expected for time series data with short-term dependencies, the ACF plot showed a distinct cutoff at lag 1, indicating a strong immediate relationship between consecutive time points. This initial spike at lag 1 in the ACF plot suggests that the ozone level at a given time is highly correlated with the ozone level in the immediately preceding time period.

After this initial spike, the ACF exhibited a gradual decay, suggesting that the influence of earlier time points persisted but weakened over time. Interestingly, the ACF dropped below 0 after a certain lag and remained negative. This negative autocorrelation indicates that higher ozone levels at a given time were followed by lower ozone levels after several lags.

The PACF plot also displayed a clear spike at lag 1, confirming the strong direct influence of the immediately preceding value. Beyond lag 1, the PACF coefficients diminished, showing weaker relationships after accounting for intermediate lags.

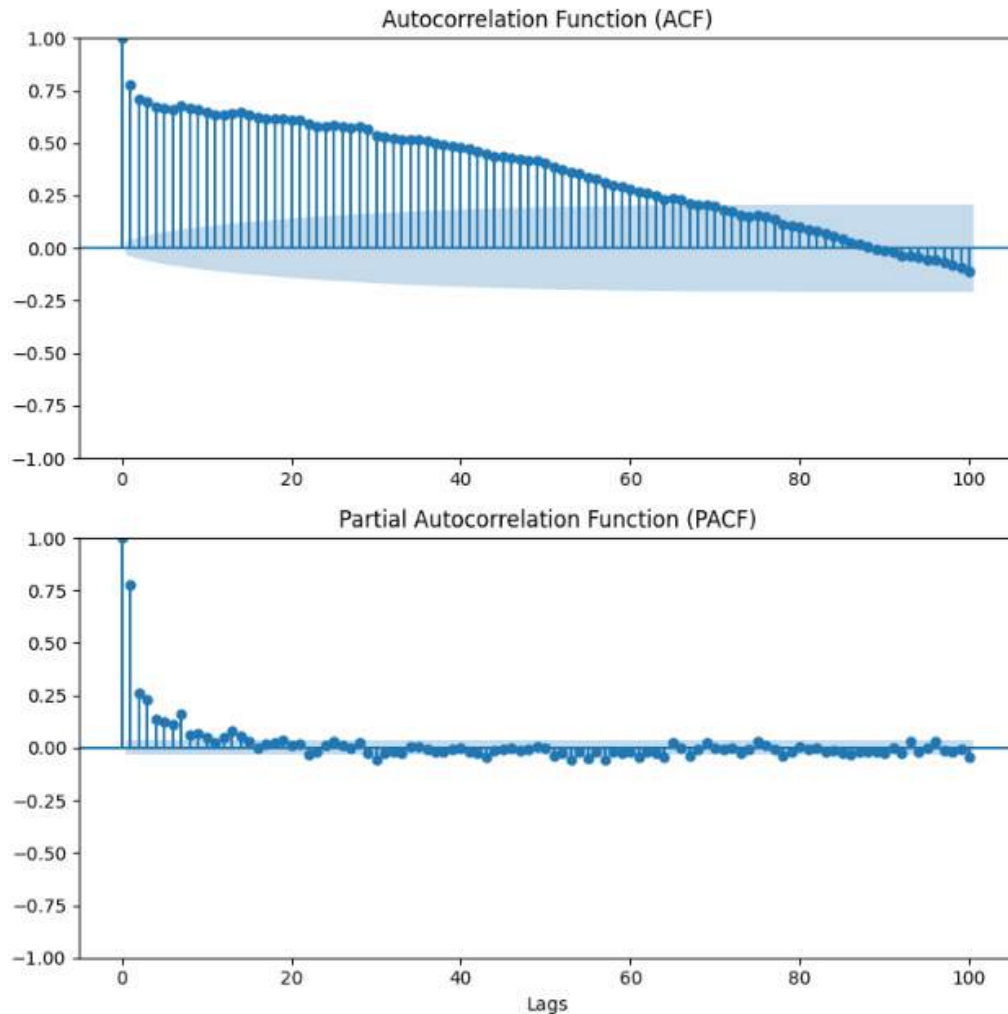


Figure 2. ACF and PACF of daily sampled ozone mean pollution time series data.

Time Series Decomposition

In order to understand trend, seasonality, and errors, we can use a technique called time series decomposition. Time series decomposition is a statistical technique used to break down a time series into its individual components: trend, seasonality, and residual. Trend is the systematic movement of the data, which can show the overall direction the data is taking over time. Seasonality is the periodic fluctuations in the data that occur at regular intervals. Residuals are the random fluctuations or noise that cannot be explained by neither the trend nor the seasonal patterns. We can use visual inspection to identify the trend of the time series as shown in Figure 3.

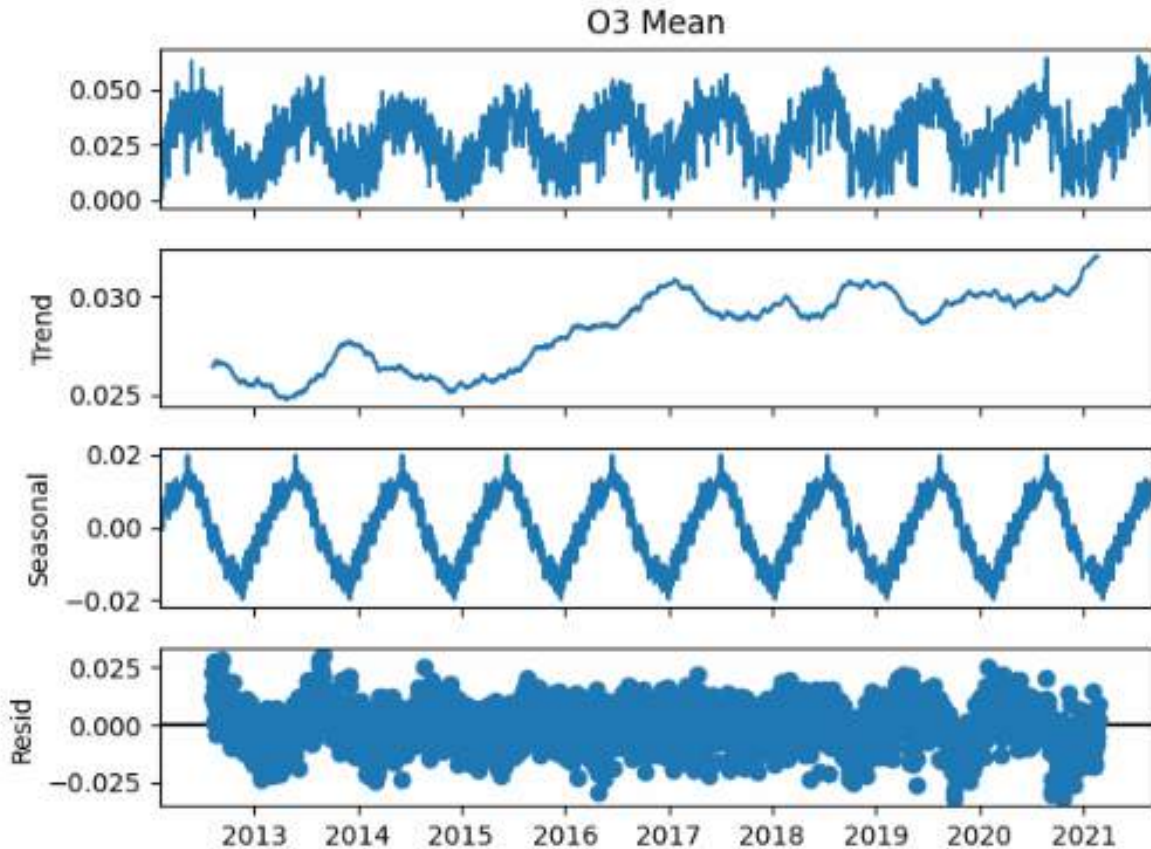


Figure 3. Time series decomposition of daily sampled ozone mean pollution data.

Data Preprocessing

In this study, we applied the rolling average to smooth the ozone level data using different window sizes: 0, 30, 60, and 90 days. The window size of 0 days represents no smoothing, showing the raw data as it is. This allows us to observe the original fluctuations and variability without any smoothing effect. The rolling average allows us to focus on long-term trends, making it easier to analyze and interpret the underlying behavior of the data. Ozone data, like many environmental time series, often contains high-frequency fluctuations that can obscure longer-term patterns. The rolling average acts as a filter, smoothing out these rapid variations and making it easier to discern the broader movements in ozone levels.

The subsequent window sizes were used to smooth out the more frequent and smaller fluctuations in the data. The choice of window sizes was made to capture different levels of long-term trends and remove short-term noise. For example, the 30-day window helps smooth out weekly or monthly fluctuations, which is useful for observing medium-term trends. By averaging over 30 days, we reduce the impact of day-to-day variability and emphasize patterns that persist for longer periods, such as monthly cycles or weather patterns. Additionally, 60-day and 90-day windows help reduce even more of the short-term noise and focus on the long-term behavior of the data, highlighting broader seasonal or yearly patterns. These longer windows

further smooth the data, allowing us to see the overall seasonal variations and multi-month trends that might be hidden in the raw data due to short-term variability.

The varying window sizes allow for a more flexible approach, with each window capturing a different range of fluctuations in the data. By comparing results with different window sizes, we can better understand how short-term vs. long-term trends influence the ozone levels over time.

Models

AutoRegressive Integrated Moving Average (ARIMA) is a statistical method used for analyzing and forecasting time series data in Python. ARIMA models are used to help understand the trends/underlying patterns in your times series data. These models can also be used to forecast future values, what we are trying to do, based on historical data. ARIMA models have three key components. The AutoRegressive (AR) part uses the past values to predict future values. The Integrated (I) part deals with non-stationary data, and uses differencing to make the data stationary. The final components, Moving Average (MA), helps identify problems or opportunities. ARIMA models are denoted as $ARIMA(p,d,q)$. The p is the order of the autoregressive part, d is the integrated part, and q is the moving average part.

In our case, after applying ARIMA to the data, we experimented with a range of values for the parameters p , d , and q based on insights from the ACF and PACF plots. Figure 2 helped us determine potential values for p and q , which relate to the autoregressive and moving average components of the model, respectively. The parameter d was chosen based on the need to make the data stationary through differencing. Based on the observed patterns in the ACF and PACF, we explored p and q values from 1 to 10 and d values also from 0 to 10, allowing for various combinations of autoregressive and moving average orders.

We tested several combinations of p , d , and q parameters and evaluated the models using their performance scores. Specifically, we used Mean Absolute Error (MAE) to measure the average difference between the forecasted and actual ozone levels. MAE is calculated by taking the absolute difference between each forecasted value and the corresponding actual value, and then averaging these absolute differences over the entire dataset. The use of absolute values ensures that both over- and under-predictions contribute equally to the error measure, without canceling each other out. A lower MAE indicates better model fit because it signifies that, on average, the model's predictions are closer to the true ozone levels. We also looked at the residuals and the forecasted values to assess how well the model fit the data. After conducting this experimentation, we identified the best ARIMA model based on a balance between the lowest evaluation score (MAE) and a reasonable fit to the data's underlying trends. The optimal model allowed us to make the most accurate forecasts of future ozone levels.

Results and Discussion

The effects of smoothing on model predictions can be both positive and negative, depending on the data and chosen smoothing technique. Positive effects include reduced noise,

improved trend identification, enhanced seasonality identification and improved model performance. Negative effects include loss of information, lagged predictions, reduced sensitivity, and spurious correlations. It is important to experiment with different smoothing techniques and window sizes to find the optimal approach for a given dataset and model. This is why we have all four types of (Figures 4-7) graphs lined one after the other to compare them. Mainly, to compare the effect of smoothing on model predictions.

The raw, unsmoothed data (Figure 4) exhibited random fluctuations, making the model sensitive to noise and short-term variations. This resulted in a less reliable prediction, as small, random changes in the data led to more erratic forecasts. This model had the highest Mean Absolute Error (MAE) of 0.0152, indicating the lowest predictive accuracy among the tested models. The higher MAE reflects the model's difficulty in capturing the underlying patterns due to the noise present in the raw data.

The 30-day smoothed data (Figure 5), with a smaller window size, helped reduce some of the noise but still remained sensitive to short-term changes. This resulted in a prediction that was smoother than the raw data but still had some small fluctuations. The MAE for this model was 0.0055, showing a significant improvement in accuracy compared to the raw data model. This demonstrates the positive impact of smoothing on reducing prediction errors by minimizing the influence of short-term noise.

The 60-day smoothed data (Figure 6) provided an even smoother prediction, with fewer short-term shifts and better representation of the underlying trends. This smoothing helped the model make more accurate predictions by reducing noise and making the trends more visible. This model achieved a slightly higher MAE of 0.0059 compared to the 30-day smoothed model. While smoothing generally improves accuracy, this slight increase suggests that there might be a point of diminishing returns, or that the 60-day smoothing might be slightly over-smoothing in this specific case.

Finally, the 90-day smoothed data (Figure 7) provided the most stable and reliable prediction. The larger window size helped the model capture long-term trends more effectively, minimizing fluctuations and improving the accuracy of future predictions. With the lowest MAE of 0.0030, this model proved to be the most accurate and reliable for forecasting ozone levels, highlighting the effectiveness of the 90-day smoothing technique in capturing long-term trends and minimizing prediction errors.

In summary, as the smoothing window increased, the predictions generally became smoother and, with the exception of the 60-day smoothing, the MAE decreased, with the 90-day smoothed data providing the best model for long-term trend forecasting. This indicates that smoothing, particularly with a larger window size, significantly improves the accuracy of ozone level predictions by reducing the impact of noise and emphasizing underlying trends. The slight increase in MAE for the 60-day smoothing suggests that optimal window size should be carefully considered, as over-smoothing might negatively impact performance.

(4, 0, 7)

MAE: 0.015163998797892106

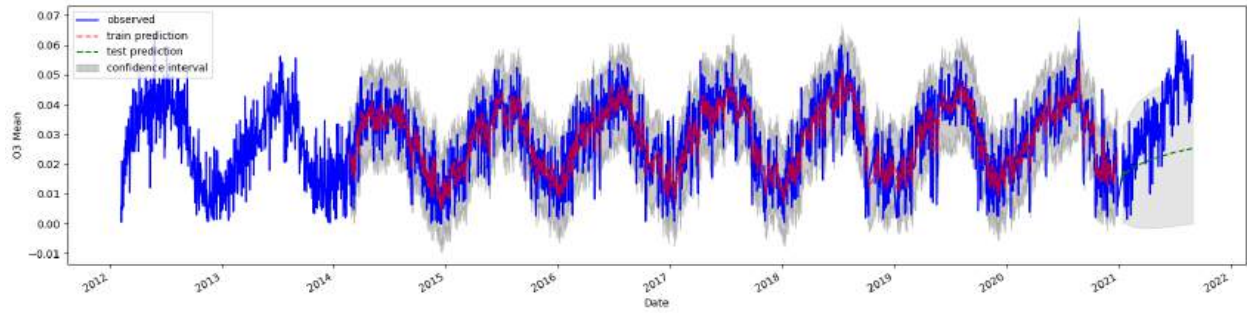


Figure 4. The best model in no day smoothing

(4, 0, 6)

MAE: 0.005536355130031676

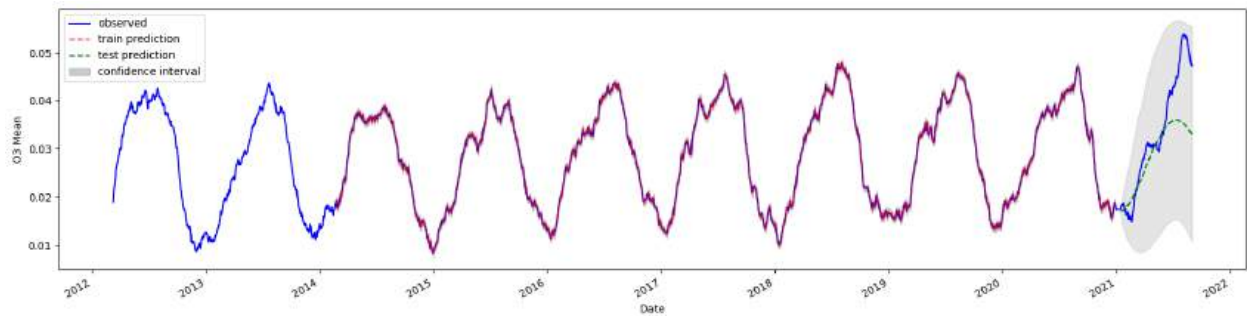


Figure 5. The best model in the 30 Day Smoothing

(4, 0, 6)

MAE: 0.005885092061579179

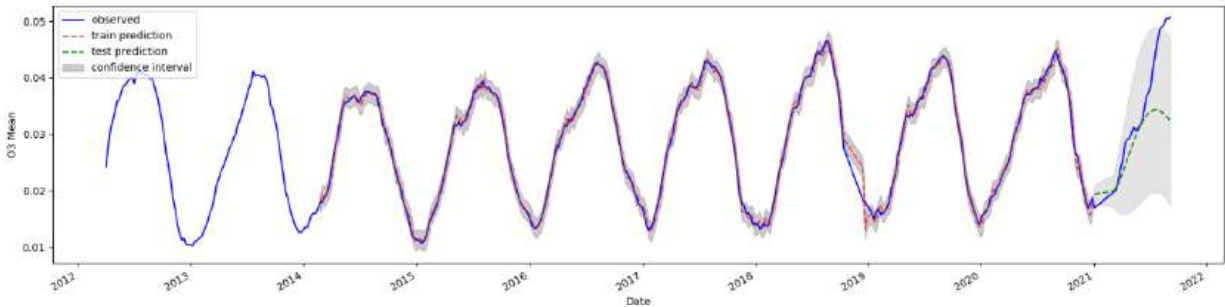


Figure 6. The best model in the 60 Day Smoothing

(4, 0, 7)

MAE: 0.002975491877552107

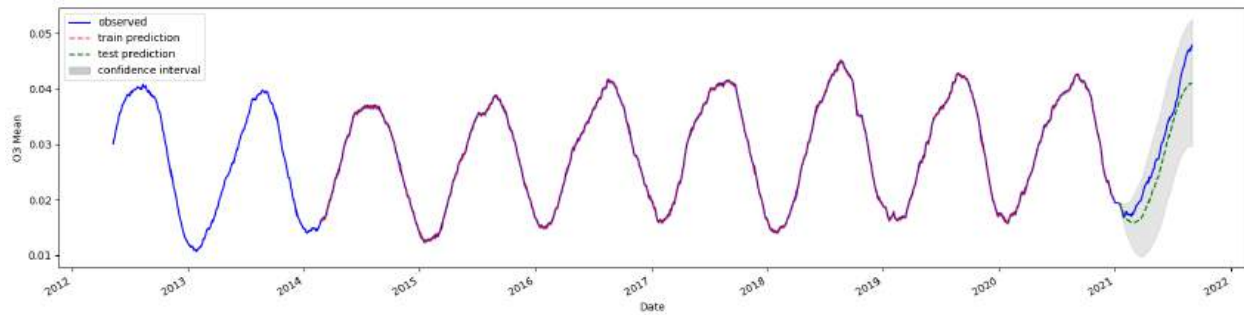


Figure 7. The best model in the 90 Day Smoothing

Conclusion

This study highlights the critical need to understand, monitor, and predict air quality trends in Colorado, particularly in Denver, where ozone pollution poses a significant threat to public health and the environment. By employing time series analysis and machine learning techniques, we gained valuable insights into the patterns and seasonality of ozone pollution, enabling more accurate forecasting. The use of smoothing techniques further demonstrated the importance of balancing noise reduction and preserving data integrity for effective model predictions.

Our findings emphasize the urgency of implementing proactive measures, such as public awareness campaigns and policy reforms, to address air quality challenges. Improved air quality would not only enhance the health and well-being of Colorado residents but also protect the state's unique ecosystems and foster economic growth by attracting tourists and businesses. The research presented here serves as a foundation for future efforts to mitigate air pollution, paving the way for a healthier and more sustainable Colorado.

In the future, we could explore the implementation of more advanced time series models, such as SARIMA (Seasonal Autoregressive Integrated Moving Average) or Prophet, which can directly handle seasonality and complex temporal dependencies without the need for pre-smoothing. These models offer greater flexibility and potentially higher accuracy by explicitly modeling the underlying data characteristics. Additionally, future research could investigate incorporating other variables, like meteorological data or traffic patterns, to improve forecasting accuracy. Furthermore, the methodologies developed in this study can be readily adapted and applied to other cities and pollutants, providing a valuable framework for air quality management and prediction on a broader scale. This approach can contribute to a more comprehensive understanding of air pollution dynamics and inform targeted interventions to improve air quality in diverse urban environments.

Improved Food Safety in Manhattan Restaurants with Data Science backed Guide to Catch Pest Violations By Rohin Vig

Abstract

Manhattan is a borough of New York City, the most populous city in North America, with 1.7 million people living there and an additional 2.3 million people commuting for work. With a population density of 17,000 people per mi² during weekdays and consuming restaurant food, public health and safety of food is paramount. This study aims to identify key factors contributing to pest infestations in Manhattan's restaurants, focusing on correlations between specific violations and restaurant characteristics to form a data-informed guide. It is built using data from a 2023 subset of the Department of Health and Mental Hygiene (DOHMH) New York City Annual Restaurant Inspection Results dataset and methods including statistical tests and machine learning models such as K-Nearest Neighbor Models, Chi-Squared Tests, Fisher's Tests, ANOVA Tests, and Multinomial Logistic Regression. This study suggests that certain types of restaurants and areas are more prone to pest issues by identifying correlations between Cuisine, Zip code, and Pest Violation Codes (Rats, Mice, Roaches, and Flies), and Other Violation Codes (Garbage not properly disposed of, Handwashing station not set up). This guide can help restaurant inspectors perform targeted inspections and restaurateurs to apply preventive measures, helping to mitigate pest-related health risks in Manhattan's densely populated areas.

Background

The dataset used in this study to identify the key factors that are correlated to pest infestations in Manhattan's restaurants is created by the DOHMH to log annual New York City restaurant citations. This dataset contains data from 2015 up till the present day and is being updated daily. 2023 data set was used for the purposes of this study as it's a complete available data set:

- 1) The mean score per citation has been increasing over the years; meaning more restaurants are having violations. The grading scale is getting tougher on the restaurants.

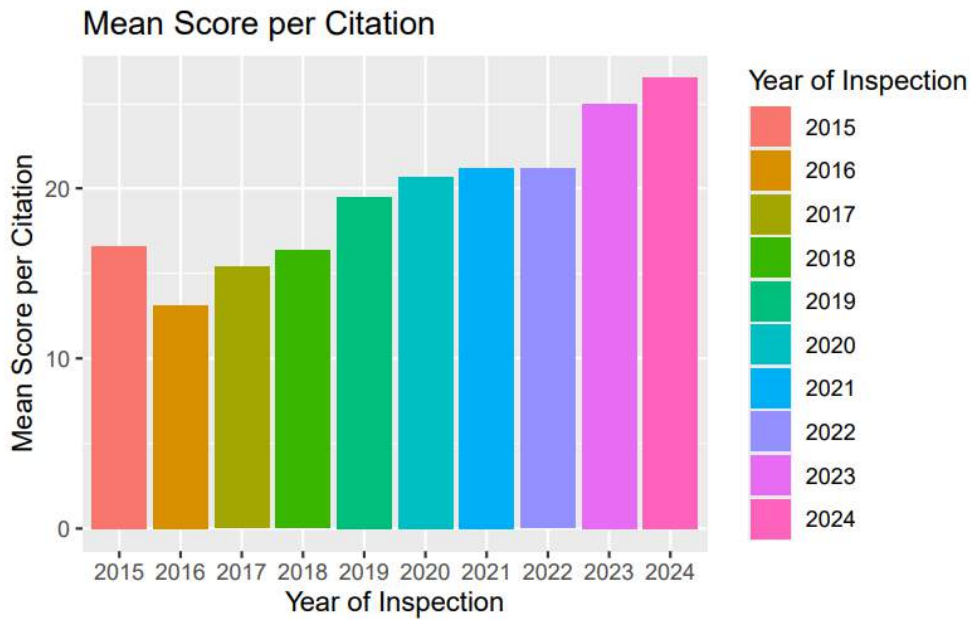


Figure 1a: The Mean Score Per Inspection is increasing over the years indicating a change in grading scale.

- 2) The number of inspections have been increasing dramatically between 2015 and 2024.

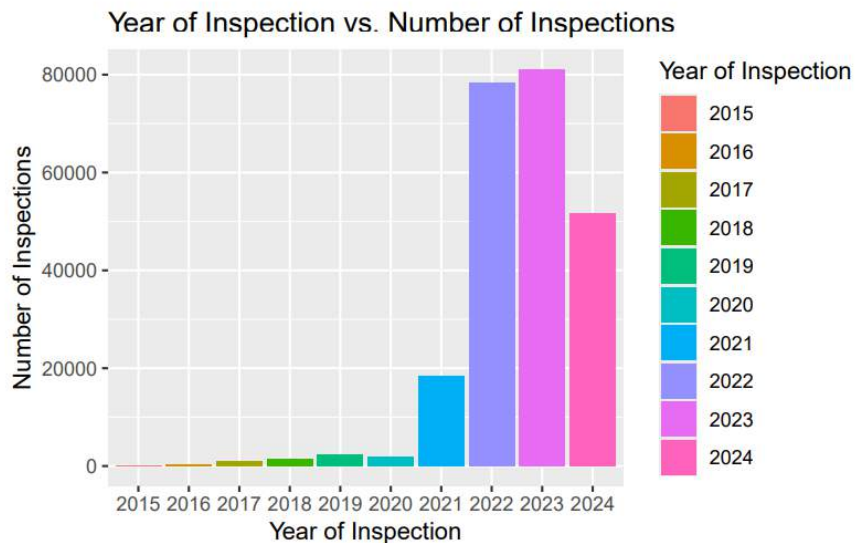


Figure 1b: The Number of Inspections is increasing over the years indicating a change in inspection policy.

METHODS

Year 2023 data set of the Department of Health and Mental Hygiene (DOHMH) New York City Annual Restaurant Inspection report is a complex data set with a lot of raw data on Manhattan Zip codes, restaurant cuisine type, type of pest violations as well as other violations. A robust software language called R was chosen for this research because R excels at performing complex statistical analyses, including linear and nonlinear modeling, time series analysis, and clustering. R is a popular choice for data scientists due to its capabilities in data manipulation, analysis, and visualization.

For this research 4 types of pest violations were considered namely Rat, Mouse, Fly and Cockroach. Few other types of violations like handwashing stations, garbage disposal were also considered. 40+ Cuisines data was considered and all Manhattan Zip Codes were analyzed.

R's statistical packages were utilized to perform Chi-Squared Tests, Fisher's test, Anova test and Multinomial Logistic Regression to find correlations between Zip Codes, Pest Violation type and Cuisine.

A **chi-squared test** (also **chi-square** or χ^2 test) is a statistical hypothesis test used in the analysis of contingency tables when the sample sizes are large. In simpler terms, this test is primarily used to examine whether two categorical variables (*two dimensions of the contingency table*) are independent in influencing the test statistic (*values within the table*).

Fisher's exact test (also **Fisher-Irwin test**) is a statistical significance test used in the analysis of contingency tables when sample sizes are small. The test is useful for categorical data that result from classifying objects in two different ways; it is used to examine the significance of the association (contingency) between the two kinds of classification.

Anova test Analysis of variance (ANOVA) is a family of statistical methods used to compare the means of two or more groups by analyzing variance. Specifically, ANOVA compares the amount of variation *between* the group means to the amount of variation *within* each group. If the between-group variation is substantially larger than the within-group variation, it suggests that the group means are likely different.

Multinomial Logistic Regression It is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables

RESULTS

Using various statistical tests and R's data visualization techniques, tests were run to find if there's any correlation among the occurrence of Pest Violation type with Cuisine and Zip Code. The results found were in direct correspondence to the hypotheses. Some of the highly dense and less affluent Zip codes were found to be high in Pest Violations. Similarly, correlations were found between Cuisine and Violation types as well as Cuisine and Zip codes as explained below.

Zip Code vs. Cuisine vs. Pest Violation Type

Figures 2a-d display a map of Manhattan with a spectrum denoting the abundance of specified pests by Zip code. These results shed light into where different populations of rats, mice, roaches, and flies reside and/or give us information on the condition of restaurants in that Zip code. Percent of Restaurants in Manhattan with Rat Infestation by Zip code



Figure 2a: Rat Violations are more common in the Harlem area and Downtown.

Percent of Restaurants in Manhattan with Mouse Infestation by Zip code

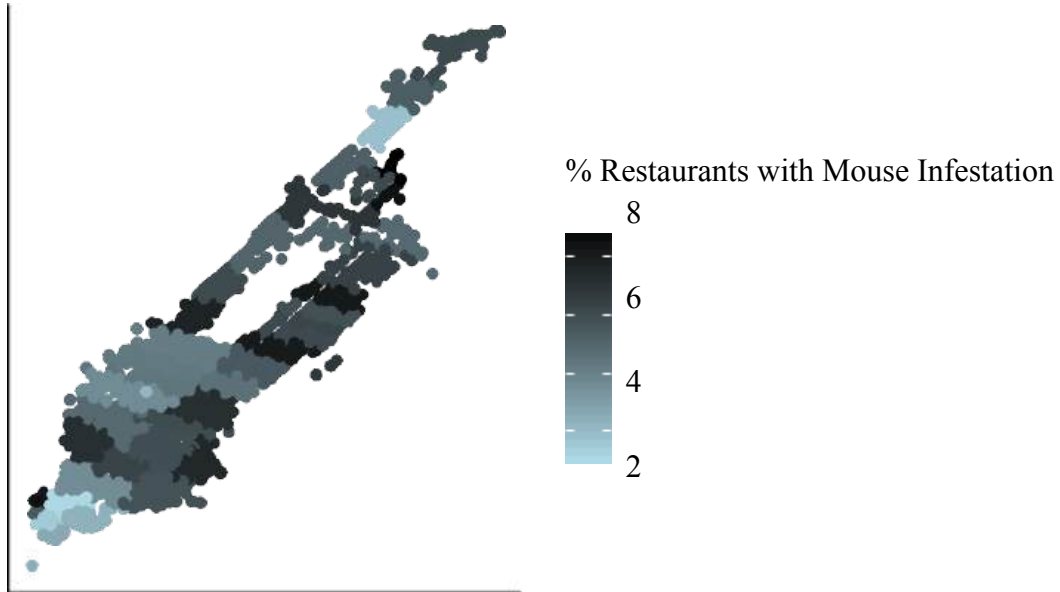


Figure 2b: *Mouse Violations are more ubiquitous than the other pests.*

Percent of Restaurants in Manhattan with Roach Infestation by Zipcode

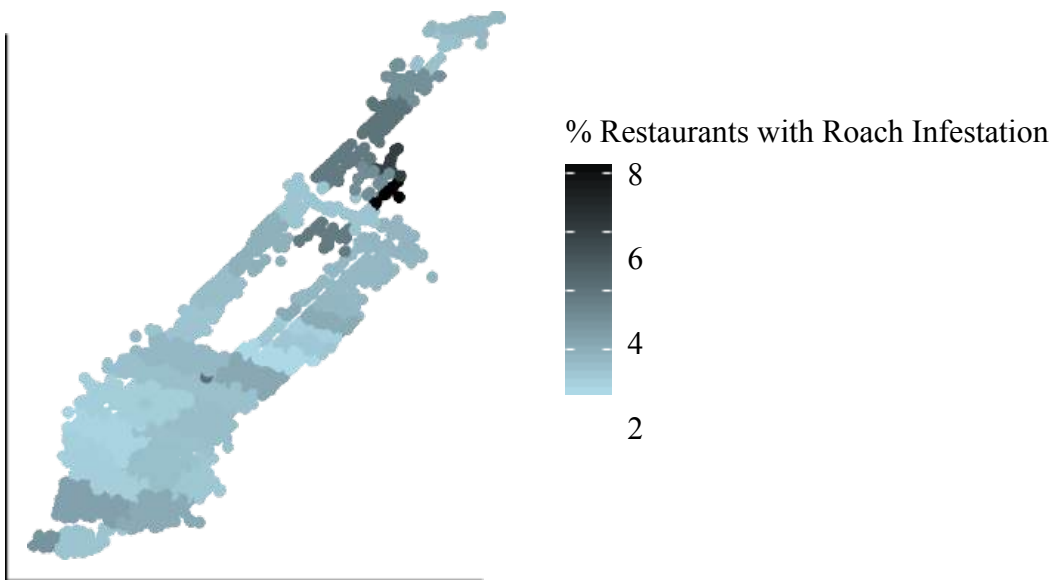


Figure 2c: *Roach Infestations are more common in the Harlem area of Manhattan.*

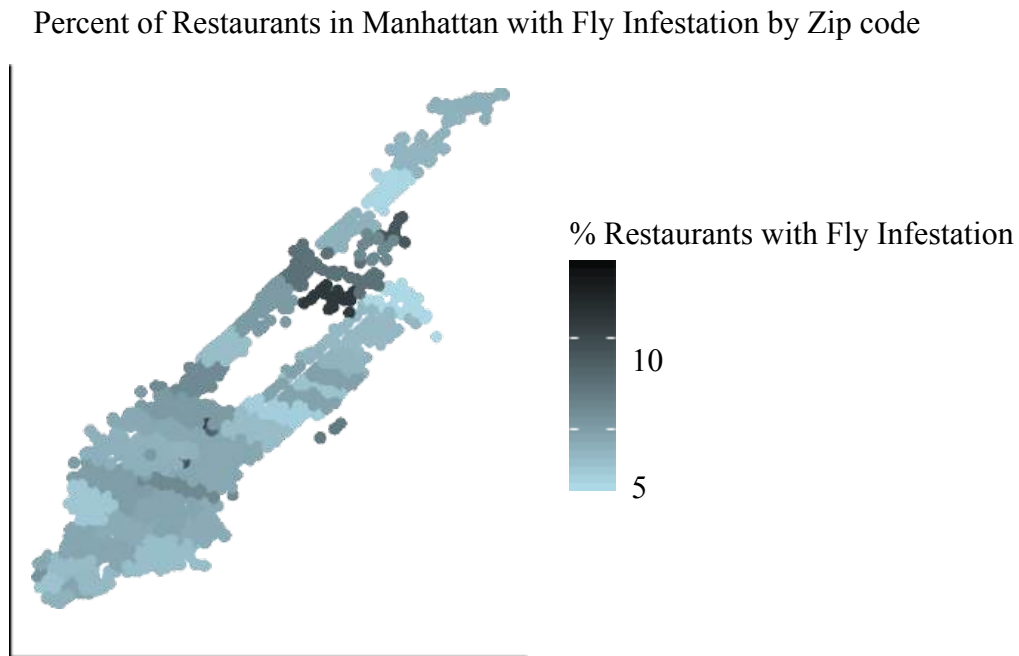


Figure 2d: *Fly Infestations are also more common in the Harlem Area.*

Figures 3a-d show which cuisines have the most percentage of restaurants with the specified pest violation by Zip code. These results give some insight into which cuisines' ingredients the pests find most attractive and/or which cuisines' restaurant establishments are most out of shape.

Cuisines with the most percentage of Rat violations by Zip code

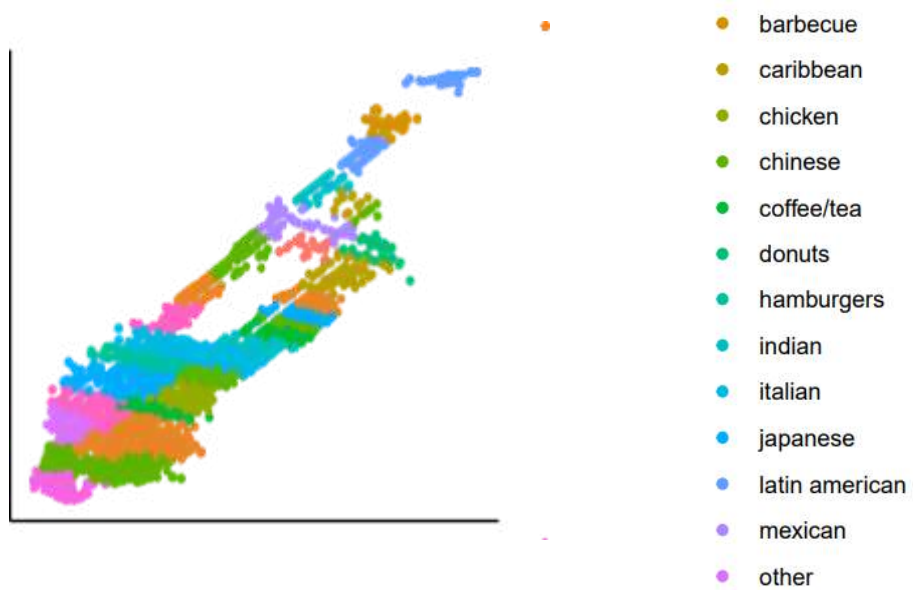


Figure 3a: Restaurant cuisines with the most percentage of Rat violations by Zip code are more varied.

Cuisines with the most percentage of Mouse violations by Zip code



Figure 3b: American cuisine is the cuisine with the most percent of Mouse Violations in Downtown New York.

Cuisines with the most percentage of Roach violations by Zip code

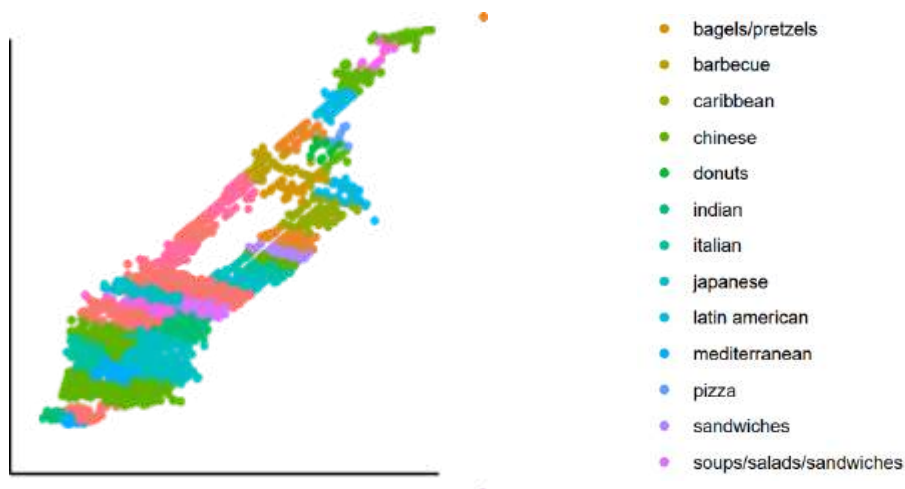


Figure 3c: Restaurant cuisines with the most percentage of Roach violations by Zip code are also very varied.

Cuisines with the most percentage of Fly violations by Zip code

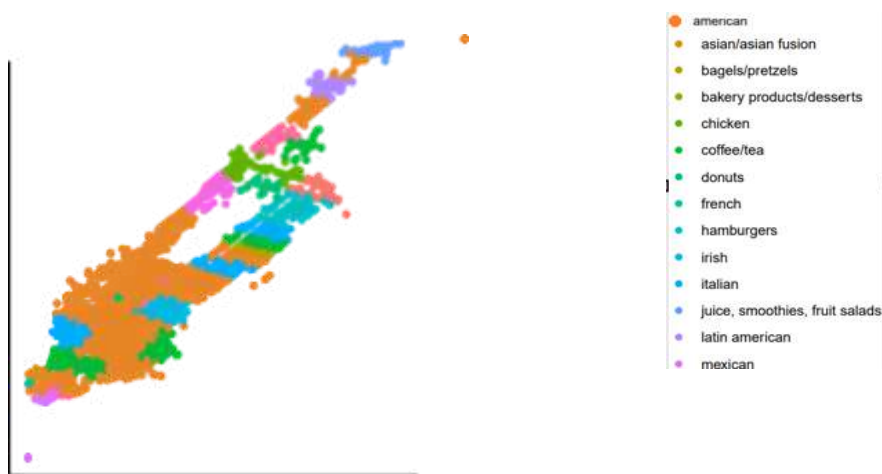


Figure 3d: American cuisine is also the cuisine with the most percent of Fly Violations in Downtown New York.

Figures 4a-d show the percent of restaurants that have a specified pest infestation by cuisine. This information can help restaurant inspectors prepare for inspections

Percent of Restaurants with Rat Violations in 2023 by Cuisine

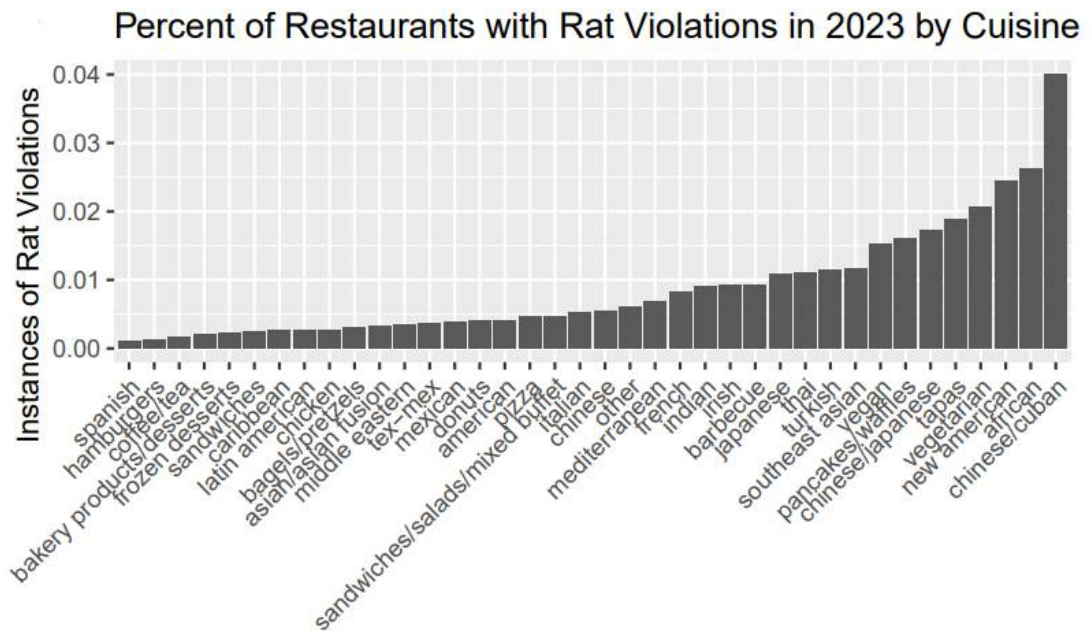


Figure 4a: Chinese/Cuban cuisine is most likely to have a Rat Violation because 4% of Chinese/Cuban cuisine has a Rat Violation.

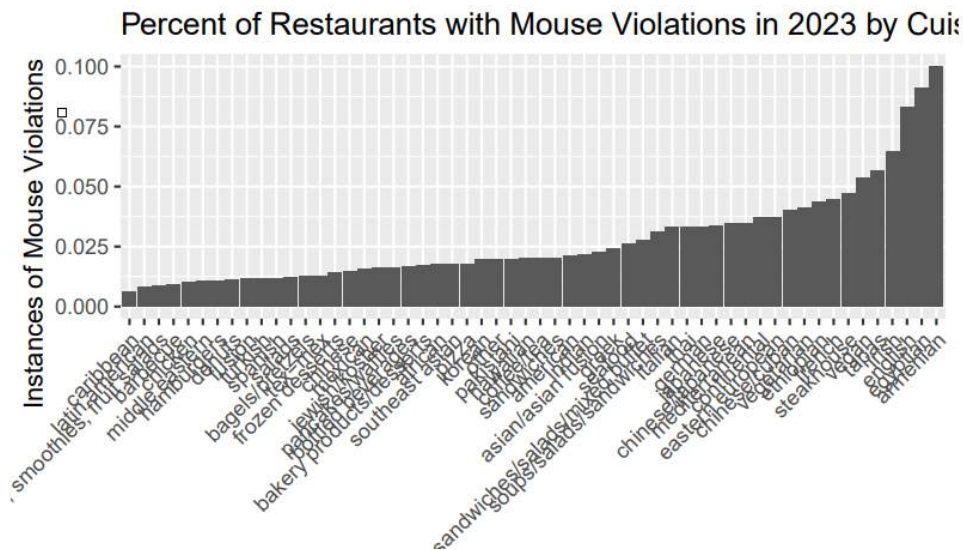


Figure 4b: Armenian cuisine is most likely to have a Rat Violation because 10% of Armenian cuisine has a Rat Violation.

Armenian cuisine has a Mouse Violation

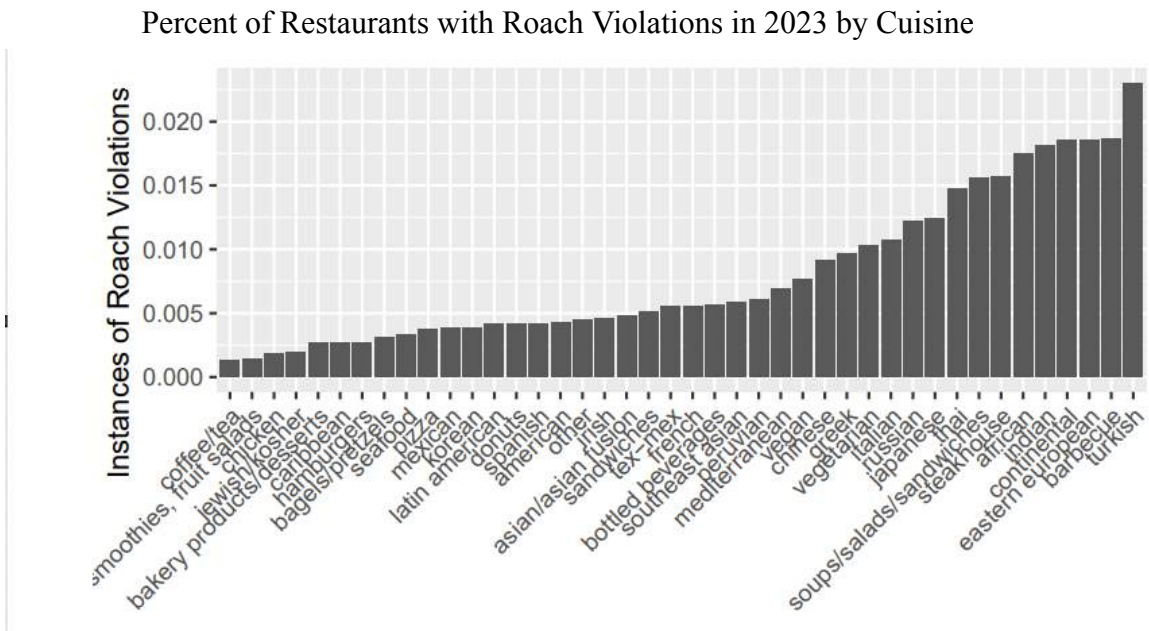


Figure 4c: Turkish Cuisine is most likely to have a Rat Violation because ~25% of Turkish cuisine has a Roach Violation.

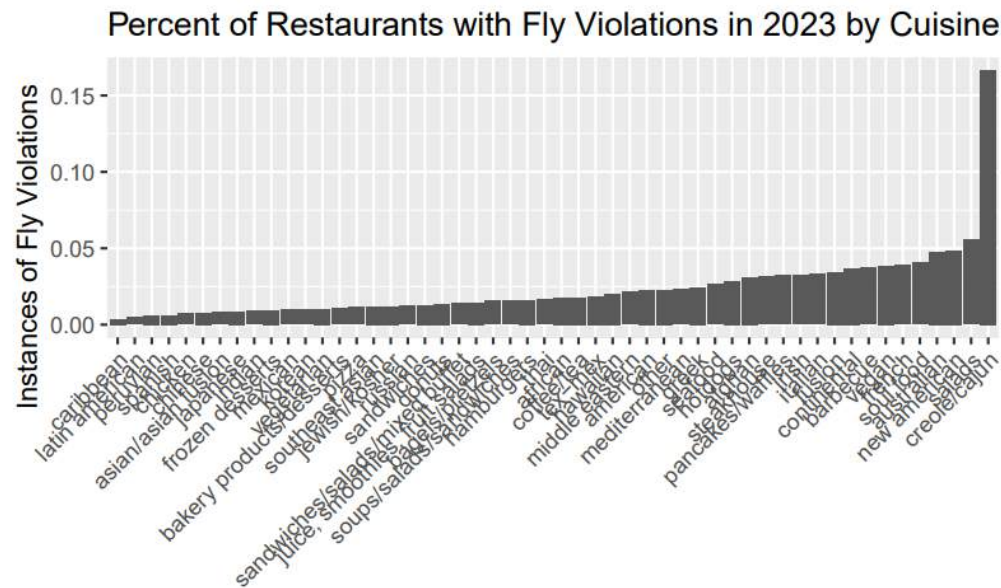


Figure 4d: *Creole/Cajun cuisine is most likely to have a Rat Violation because ~17% of Creole/Cajun cuisine has a Fly Violation.*

Statistical Tests

Cuisine Vs Pest Violation Type : Multinomial Logistic Regression was used to find the correlations between Pest Violation Type and Cuisine. This model is used to model outcomes where the dependent variable is categorical and has more than many levels. Observation is that there are certain Pest Violation Types and Cuisines that are correlated with each other, but not all of them are correlated. Below table lists the ones that are correlated:

Cuisine	Violation Type	P-Values
Asian/Asian Fusion	Fly	0.016
Australian	Fly	0.018
Bakery Products/Desserts	Mouse	0.045
Barbecue	Mouse	0.035
Coffee/Tea	Mouse	1.026e-05
Chicken	Mouse	0.015
Creole/Cajun	Fly	0.019
French	Roach	0.004
Fusion	Fly	0.003
Hamburgers	Fly	0.010
Indian	Roach	0.026
Latin American	Mouse & Fly	0.011, 0.029
Middle Eastern	Mouse	0.022
Peruvian	Mouse	0.031
Salads	Mouse	1.021e-14
Sandwiches	Mouse	0.006
Soul Food	Mouse	0.046
Soups/Salads/Sandwiches	Mouse	0.034
Steakhouse	Fly	0.00
Thai	Rat	2.597e-12

Figure 5a: *The specified restaurant is **more likely** to have the specified pest violation than any other. P-values are under = 0.05, therefore Violation Type and Cuisine have a statistically significant correlation.*

Cuisine vs. Zipcode Correlation

A Chi-Squared Test as well as a Fisher's Test was used to test the correlation between Cuisine and Zip code. These tests were used because they compare two categorical variables, and Cuisine and Zip code fall under that description. I found that there is indeed a correlation between Cuisine and Zip code. This means that Figures 3a-d are backed by statistical significance.

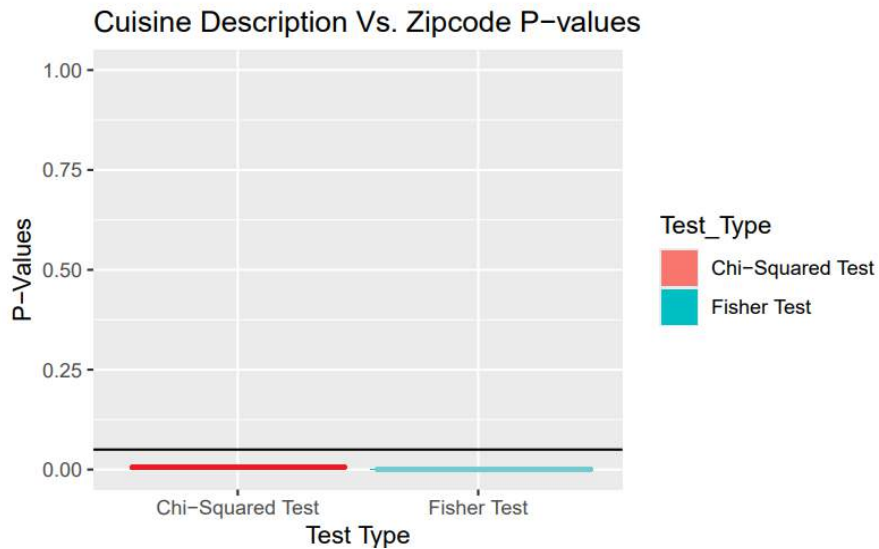


Figure 5b: *P-values are under $= 0.05$, therefore Zipcode and Cuisine have a statistically significant correlation*

Pest Violation Type vs. Zip code and Cuisine Correlation

An Anova Test was used to test the correlation between Zipcode & Cuisine and Pest Violation Code. It was used because Anova tests are used to test the correlation between a categorical numerical variable as well as a categorical variable. Pest Violation Type was transformed into a numerical categorical variable by calculating the average number of each Pest Violation Type grouped by Cuisine and Zip code for each Anova test. I found that there is a correlation between both Cuisine and Pest Violation Type & Zip code and Pest Violation Type. This also means that Figures 2a-d and Figures 4a-d are backed by statistical significance.

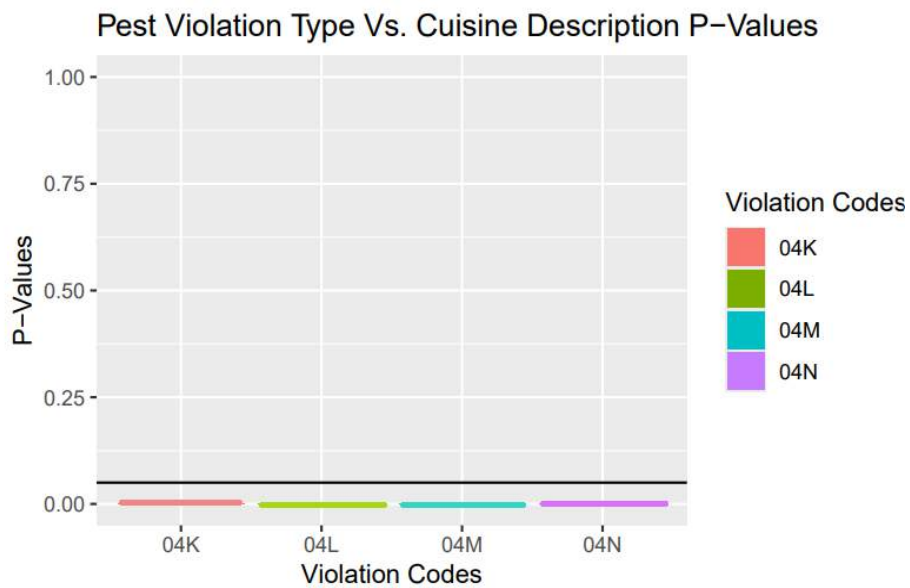


Figure 5c: The correlation between Violation Code and Zipcode is Statistically Significant

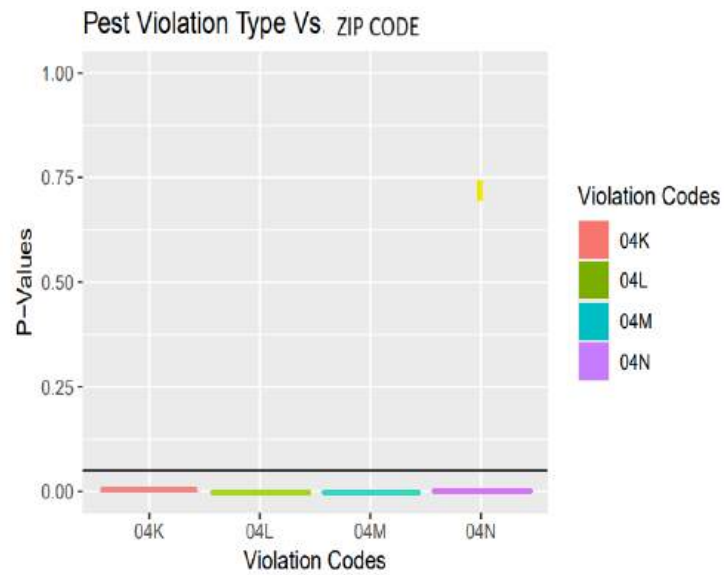


Figure 5d: The correlation between Pest Violation Type and Zip code is Statistically Significant

Pest Violation Type vs. Other Violation Type Correlation

Violation	Pest Violations	P-Values
Food not protected from cross-contamination	Rats, Roaches, Flies	0.005091, 0.008444, 6.159e-05
Culinary sink or alternative method not provided for washing food	Rats	0.0007472
Hand washing facilities not provided or not located where required	Rats, Roaches, Flies	0.01913, 0.0007209, 0.02828
Food not properly protected when stored	Roaches	0.003129
Equipment with food contact surfaces not clean and sanitized	Mice	0.003348
Prevention and control measures not used for pest management	Rats, Mice, Roaches, Flies	< 2.2e-16, < 2.2e-16, < 2.2e-16, < 2.2e-16
Garbage not properly removed or stored	Mice, Flies	0.0198, 3.376e-13
Pesticides not properly labeled, not authorized for use, or improperly used	Rats, Mice, Roaches, Flies	2.137e-10, 9.836e-14, 2.202e-11, 7.726e-05
Toilet facility not properly maintained	Rats	0.00242
Improper disposal of sewage or liquid waste	Mice, Flies	0.03035, 0.006392
Thermometers not provided in cold/hot storage	Roaches, Flies	0.006546, 0.003033
Non-food contact surfaces (wall, ceiling, floors) improperly constructed/maintained	Mice, Flies	8.566e-07, 9.552e-05

Figure 6: The table above shows which Other Violation Types that a restaurant has that makes it more likely to harbor Rats, Mice, Roaches, or Flies.

Discussion

Using statistical tools, correlations were found between Zip code, Cuisine and Pest violation type. These results can prove useful in restaurant inspection as a guide to health inspectors. Firstly, health inspectors can look up the Zip code and cuisine on the guide and see which violations similar restaurants are likely to have and prepare for their visit. Secondly, during the inspection, health inspectors can look up which pest violation a restaurant is likely to have from easily discernible characteristics from the kitchen like whether or not a proper waste disposal system is in place.

Some future work that would be very useful in making health inspection and restaurant upkeep easier would be to investigate whether or not there is any causation between environmental factors in New York Streets and pest abundances in that zipcode. It would also be useful to do the same analysis with the 2024/2025 dataset for an up-to-date guide.

Works Cited

- The data is from the DOHMH's NYC OpenData website: https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/43nn-pn8j/about_data
- https://en.wikipedia.org/wiki/Chi-squared_test
- https://en.wikipedia.org/wiki/Fisher%27s_exact_test
- https://en.wikipedia.org/wiki/Multinomial_logistic_regression#cite_note-1
- https://en.wikipedia.org/wiki/Analysis_of_variance
- Fisher, R. A. (1922). "On the interpretation of χ^2 from contingency tables, and the calculation of P". *Journal of the Royal Statistical Society*. **85** (1): 87-94. doi:10.2307/2340521. JSTOR 2340521.
- Fisher, R.A. (1954). *Statistical Methods for Research Workers*. Oliver and Boyd. ISBN 0-05-002170-2.
- "Chi-Square - Sociology 3112 - Department of Sociology - The University of Utah". *soc.utah.edu*. Retrieved 2022-11-12.
- Greene, William H. (2012). *Econometric Analysis* (Seventh ed.). Boston: Pearson Education. pp. 803–806. ISBN 978-0-273-75356-8.
- R Tutorials:
- <https://www.w3schools.com/r/>
- <https://www.youtube.com/watch?app=desktop&v=9fUKihCsDag&t=18>
- <https://www.codecademy.com/learn/learn-r>

The Imperative of Ethical Leadership: Lessons from the Mahabharata on How King Dhritarashtra's Bias, Poor Counsel, and Selfish Governance Led to Political Instability, Economic Decline, and the Suffering of His People – Parallels with Modern Political and Business Failures By Varun Ganesh

Abstract

The Mahabharata is not just an Itihasa; it is a profound meditation and discourse on righteousness, leadership, and the consequences of moral blindness. King Dhritarashtra of Hastinapur, physically blind, symbolized the dangers of moral and political blindness. His inability to see the destructive nature of his biased decisions, influenced by his attachment to his sons, led to the catastrophic Kurukshetra War, tearing apart his kingdom and causing unimaginable suffering. Lord Krishna's counsel in the *Bhagavad Gita*—"न हि कश्चित्क्षणमपि जातु तिष्ठत्यकर्मकृत्" (*Na hi kaścit kṣaṇam api jātu tiṣṭhaty akarmakṛt*, Gita 3.5)—underscores that no one can escape the consequences of their actions, emphasizing the need for conscious leadership.

This paper dives into Dhritarashtra's flawed leadership, exploring how personal biases, poor counsel, and unchecked loyalty can bring about destruction. Drawing comparisons to modern leaders—such as Robert Mugabe and Jeffrey Skilling—this paper shows that the core lessons from Dhritarashtra's story are still relevant today. By examining these ancient examples alongside modern case studies, we can gain insights into the importance of impartiality, sound decision-making, and the need for leadership that serves the common good, not personal interests. As the *Bhagavad Gita* advises, "When meditation is mastered, the mind is unwavering like the flame of a lamp in a windless place" (*Bhagavad Gita* 6.19), underscoring the need for clarity and detachment in decision-making. Additionally, *The Leadership Challenge* by Kouzes and Posner stresses the importance of leaders who can "model the way" and make decisions based on values rather than personal bias, a critical lesson we can learn from both the Mahabharata and modern leadership failures. Furthermore, Sri Aurobindo, in his "Essays on the Gita," presents a synthesis of love, knowledge, and action as pathways to spiritual and ethical leadership. Additionally, "The Long Walk of Peace," published by UNESCO, provides insights into conflict prevention and resolution, emphasizing the role of cultural understanding and proactive measures in fostering peace. These perspectives collectively reinforce the imperative of ethical leadership in both political and business realms. By analyzing King Dhritarashtra's failures and juxtaposing them with global case studies, this paper highlights the urgency of ethical leadership, impartial decision-making, and sound governance.

Introduction

The story of King Dhritarashtra, a ruler metaphorically blind to the principles of *dharma*, serves as a timeless cautionary tale. His deep attachment to his sons, particularly Duryodhana, led him to make biased and irrational decisions, ultimately causing the downfall of his kingdom and immense suffering for his people. Despite having access to wise advisors like Vidura and

Bhishma, who repeatedly urged him to act justly, Dhritarashtra's emotional entanglement clouded his judgment. Instead of ruling with righteousness, he allowed his personal attachments to dictate his governance, demonstrating a pitfall that leaders across history have struggled with. Vidura, one of Dhritarashtra's most trusted advisors, warned him of the consequences of ruling with bias and neglecting justice. As Vidura cautions in the *Udyogaparvan*: "न ते राज्यं स्वमित्युक्तं प्रजासंविभजं नृप। धर्मेण तस्य संरक्षः प्रजानां च हिते रतः॥" (*Na te rājyaṁ svam ityuktam prajāsaṁvibhajaṁ nṛpa, Dharmeṇa tasya saṁrakṣaḥ prajānāṁ ca hite rataḥ.*)—*Mahabharata, Udyogaparvan 5.13* ("The kingdom is not a personal possession, O King; it must be shared with the people. It should be protected with righteousness, always keeping the welfare of the subjects in mind.")

Yet, Dhritarashtra failed to heed these words. His inability to rule impartially not only led to war but also destabilized Hastinapur's economic and political systems. Vidura's wisdom extended further, reminding him: "राजा धर्मेण संयुक्तः प्रजानां परिपालकः। स स्वर्गं याति धर्मात्मा सर्वभूतहिते रतः॥" (*Rājā dharmeṇa saṁyuktaḥ prajānāṁ paripālakaḥ, Sa svargaṁ yāti dharmātmā sarvabhūtahite rataḥ.*)—*Mahabharata, Udyogaparvan 5.13* ("A king who upholds righteousness and protects his subjects shall ascend to heaven, for he is devoted to the welfare of all beings.") Dhritarashtra's continued disregard for wise counsel eventually led to catastrophe. His poor parenting, acceptance of wicked schemes against the Pandavas, and reliance on the deceitful strategies of his brother-in-law Shakuni all contributed to Hastinapur's downfall. As Bhishma warns in the *Sabhaparvan*: "अन्यायेन तु यो राजा राज्यं इच्छति धारयुम्। स शत्रुभिर्हर्यविध्वस्तः शीघ्रं विनश्यति स्वयम्॥" (*Anyāyena tu yo rājā rājyaṁ icchati dhārayum, Sa śatrubhirhyavidhvastaḥ śīghraṁ vinaśyati svayam.*)—*Mahabharata, Sabhaparvan 3.10* ("A king who lets his kingdom slip into the hands of the unjust brings ruin upon his people and is swiftly destroyed by his enemies."). The collapse of his lineage and kingdom serves as a stark reminder of the far-reaching consequences of unethical leadership.

This paper explores King Dhritarashtra's decisions and their profound impact on Hastinapur, drawing comparisons to modern leadership, where bias, poor counsel, and self-interest have similarly led to crises. Scholars like Emily Hudson have extensively analyzed the *Mahabharata*'s themes of moral blindness and suffering, particularly in *Disorienting Dharma: Ethics and the End of the World in the Mahabharata* (Hudson, 2014). She argues that "Moral blindness in the Mahabharata is not merely an individual failing but a systemic one, where personal biases and ethical lapses on the part of those in power result in widespread suffering" (Hudson, 2014, p. 82). This is evident in Dhritarashtra's leadership, where his failure to prioritize the greater good led to widespread devastation.

Similar patterns emerge in modern political history. The Watergate scandal (1972-1974) in the United States exposed how Richard Nixon's prioritization of personal survival over institutional integrity led to a constitutional crisis. Likewise, Argentina's hyperinflation in the late 20th century, fueled by short-term economic policies, echoes the economic downfall caused by Dhritarashtra's poor decision-making. The *Bhagavad Gita* reinforces the importance of detached and duty-driven leadership: "कर्मण्येवाधिकारस्ते मा फलेषु कदाचन।"

(*Karmaṇy-evādhikāras te mā phaleṣu kadācana.*)—*Bhagavad Gita* 2.47. ("Your right is to perform your duty alone, never to its fruits.")

In both political and business leadership today, similar challenges arise. When leaders prioritize personal loyalties, greed, or short-term gains over the collective good, they risk destabilizing entire nations and corporations. The *Mahabharata* teaches the necessity of impartial governance, ethical decision-making, and the prioritization of long-term stability over immediate personal gains. As Hudson's analysis underscores, leaders must rise above their personal attachments to act justly. Lord Krishna's teachings in the *Bhagavad Gita* (Gita 3.35) reinforce this, stating that "*A wise person does not act from personal desire but according to the needs of the time and place.*"

By examining both ancient wisdom and modern leadership failures, the lessons from the *Mahabharata* continue to hold relevance today, offering valuable insights into the necessity of ethical governance in politics and business.

This paper will further explore how Dhritarashtra's leadership failures parallel contemporary political and business crises, highlighting the essential role of ethical leadership in ensuring stability and prosperity.

Dhritarashtra's Bias and Its Impact on Political Stability and Economics

King Dhritarashtra's emotional attachment to his sons, particularly Duryodhana, played a central role in his blind decisions that ultimately led to the Kurukshetra War and the downfall of Hastinapur. Despite the warnings of his wise advisors, such as Vidura and Bhishma, Dhritarashtra allowed his loyalty to his family to override his duty to his people. His actions exemplify the dangers of excessive attachment, as emphasized in the *Bhagavad Gita*. Lord Krishna cautions about the perils of emotional entanglement and how it clouds judgment: "इन्द्रियाणां हि चरतां यन्मनोऽनुविधीयते। तदस्य हरति प्रज्ञां वायुर्नवमिवाम्भसि॥" (*Indriyāṇāṃ hi caratām yan mano 'nuvidhīyate, Tadasya harati prajñām vāyur nāvam ivāmbhasi.*)—*Bhagavad Gita* 2.67 ("Just as a boat on water is swept away by a strong wind, the mind that follows wandering senses loses its wisdom.")

This verse underscores the importance of emotional detachment in leadership, stressing that attachment can impair a leader's ability to make rational, objective decisions. Dhritarashtra's inability to detach from his sons' actions, despite their moral failings, mirrors the behavior of contemporary leaders who, driven by personal biases, destabilize their nations. For example, Robert Mugabe's rule in Zimbabwe was marked by a similar attachment to a corrupt inner circle, ultimately leading to the economic collapse and political instability in the country. Mugabe's loyalty to his circle of power and his disregard for the broader public good led to disastrous consequences for his people, much as Dhritarashtra's attachment to his sons caused suffering for his kingdom.

From a management perspective, emotional attachment to personal interests or individuals can cloud a leader's judgment and impede effective decision-making. James Kouzes and Barry Posner, in their book *The Leadership Challenge*, argue that leaders who fail to act with

objectivity and prioritize the best interests of their organization often lose credibility and trust. They emphasize that leadership is grounded in the ability to make decisions for the collective good, even when those decisions may be difficult or unpopular. Kouzes and Posner state: "*Leaders who are not objective and fail to act in the best interest of the organization erode trust and credibility, which undermines their effectiveness.*" (Kouzes & Posner, 2017, p. 85) Dhritarashtra's leadership exemplifies this principle—his inability to make impartial decisions based on sound advice led not only to the fall of his family but also to the destruction of his kingdom.

Furthermore, Dhritarashtra's leadership failure offers valuable insights into the broader implications of attachment in both politics and management. In the world of business, attachment to personal ambitions or vested interests can similarly lead to poor governance. A case in point is the collapse of Enron, where CEO Jeffrey Skilling's personal greed and attachment to an unethical corporate culture resulted in one of the largest corporate collapses in history. Skilling, like Dhritarashtra, ignored warnings from advisors and focused instead on maintaining personal power, ultimately causing massive financial and reputational damage to the company. Similarly, recent failures such as the downfall of WeWork's founder Adam Neumann (2019) and the FTX cryptocurrency collapse (2022) illustrate how unchecked ambition and emotional decision-making can lead to disaster, much like Dhritarashtra's failure to control his attachments. By drawing from ancient wisdom and contemporary examples, the *Mahabharata* provides a compelling case for leaders to practice detachment, rational governance, and a commitment to the greater good.

The Economic and Social Costs of Poor Leadership

The aftermath of the Kurukshetra War was nothing short of devastating for Hastinapur. The destruction of its resources, the loss of skilled workers, and the collapse of the economy left the kingdom in ruin. The once-thriving land was reduced to poverty, with its citizens living in despair. Much like the destruction of modern economies in the wake of war, Hastinapur's devastation was far-reaching, affecting not only the current generation but also those that would come after. The war also meant the loss of the entire male population of the kingdom, sons, fathers, and brothers, all gone. This left families without their primary breadwinners and role models. The social fabric of Hastinapur was torn apart, and the next generation grew up without the guidance and support of men in the family. This cycle of loss, both emotional and material, created long-lasting effects that stunted the kingdom's recovery and growth.

This scenario in the *Mahabharata* highlights the long-term costs of poor leadership decisions, especially when leaders are blinded by personal attachments and fail to act in the best interests of their people. The economic devastation of Hastinapur can be compared to the economic collapse caused by wars and mismanagement in the modern world. Similar tragedies are evident in nations like Lebanon, where financial mismanagement and political corruption led to an economic crisis in 2019, and in Sri Lanka's 2022 economic collapse, where poor leadership decisions resulted in mass protests and financial turmoil.

Lord Krishna in the *Bhagavad Gita* explains the profound impact of a leader's moral choices on the collective destiny of their people: "द्वाविमौ पुरुषौ लोके क्षरश्चाक्षर एव च।क्षरः सर्वाणि भूतानि कूटस्थोऽक्षर उच्यते॥"(Dvāv imau puruṣau loke kṣaraś cākṣara eva ca, Kṣaraḥ sarvāṇi bhūtāni kūṭastho 'kṣara ucyate.)—*Bhagavad Gita* 15.16. ("There are two types of beings in this world—the perishable and the imperishable. All living entities are perishable, but the one who remains steadfast in righteousness is imperishable.")

Dhritarashtra's failure to choose the path of righteousness led not only to the downfall of his kingdom but also to the emotional and economic devastation of his people. His poor leadership exacerbated the suffering of his citizens and led to irreversible social consequences. Just as Dhritarashtra's attachment to his son clouded his judgment, modern leaders who prioritize short-term political gains over long-term stability create crises that extend far beyond their tenure.

From a management perspective, the economic and social consequences of poor leadership are well-documented in business and organizational theory. In their seminal work, *The Leadership Challenge*, Kouzes and Posner (2017) argue that a leader's failure to act with integrity and focus on the common good can lead to significant damage to the organization. They stress that effective leaders must ensure the health and growth of their organization, not just in the short term but for the long haul:

"The most enduring leadership challenges are those that require long-term solutions, and leaders must be willing to make decisions that benefit the long-term health of the organization, even when those decisions are unpopular." (Kouzes & Posner, 2017, p. 98)

In the modern business world, we see similar patterns of destruction. The Boeing 737 MAX crisis (2018-2019) resulted from prioritizing profit over safety, leading to tragic accidents, loss of public trust, and billions in financial losses. Additionally, the downfall of Theranos (2015-2018), driven by Elizabeth Holmes' deceptive leadership and disregard for ethical business practices, mirrors Dhritarashtra's failure to uphold *dharma*, resulting in the collapse of an empire built on illusions.

The lessons from such failures parallel Dhritarashtra's leadership mistakes: failing to make ethical decisions for the greater good of the community ultimately harms everyone involved. By drawing from ancient wisdom and present-day examples, the *Mahabharata* provides a compelling case for leaders to act with foresight, ethical integrity, and responsibility toward their people.

The Impact of War on People's Happiness, Business, and Growth

War is not only physically destructive; it also leaves a deep emotional scar on a population. The happiness of the people is crushed as they lose their loved ones, their homes, and their livelihoods. In Hastinapur, once a prosperous kingdom, the effects of the Kurukshetra War were catastrophic. Citizens who had once lived comfortably in peace were suddenly thrust into

an environment of grief, loss, and deprivation. The kingdom's businesses were decimated, its economy shattered, and the societal structures that had once supported the people were torn apart. In the wake of the war, the youth of the kingdom grew up in an environment of despair, robbed of the guidance and mentorship of their fathers and male role models.

As the *Striparvan* of the *Mahabharata* poignantly illustrates, the human cost of war is incalculable. The devastation is not merely in the physical destruction of land and property, but in the emotional and psychological toll it exacts on the people who survive. In this book, James Fitzgerald's translation vividly captures the overwhelming grief that war brings to families and the society at large. In the *Striparvan*, the survivors mourn the loss of their loved ones, acknowledging that the battlefield's victory came at the ultimate cost—human lives and the mental and emotional well-being of the populace (Fitzgerald, 2004). This criticism of warfare, highlighting the immense suffering it brings, serves as a warning against the perils of unchecked aggression and the price of war, which extends far beyond the battlefield.

Lord Krishna's teachings in the *Bhagavad Gita* underscore the destructive effects of war on personal well-being and societal growth. Krishna advises Arjuna: "य एनं वेति हन्तारं यश्चैनं मन्यते हतम्। उभौ तौ न विजानीतो नायं हन्ति न हन्यते॥" (*Ya enam veti hantāraṁ yaś cainam manyate hatam, Ubhau tau na vijānīto nāyam hanti na hanyate.*)—*Bhagavad Gita* 2.19 ("He who thinks the self slays or is slain does not understand. The self slays not, nor is it slain.")

This verse speaks to the futility of attachment to outcomes, particularly in the context of war. In the aftermath of the Kurukshetra War, the people of Hastinapur were left to confront the harsh reality of their shattered society. The death of the warriors and the physical and emotional toll of the war stifled any potential for future growth. The constant cycle of grief and despair stunted the possibility of healing, let alone economic recovery. People were no longer able to take joy in their daily work, and the kingdom's once-thriving businesses were destroyed, unable to recover from the devastation.

This emotional and economic collapse mirrors what often happens after wars in modern contexts. For example, post-conflict nations such as Ukraine, following the 2022 invasion, have faced long-term economic stagnation and psychological trauma. The destruction of infrastructure, displacement of millions, and prolonged instability have led to a dramatic decline in GDP and employment, much like the struggles of the people of Hastinapur in the *Mahabharata*.

From a business perspective, the long-term consequences of war on economic growth and stability have been well-documented. As Schilling (2020) discusses in her book *Strategic Management: Theory and Cases*, a disruption of business continuity, whether from war or economic collapse, results in long-term stagnation. Companies often face severe operational setbacks, loss of human capital, and the destruction of infrastructure. The *Mahabharata* highlights that, just as the war destroyed Hastinapur's infrastructure and human resources, modern economies similarly suffer when wars lead to the loss of skilled workers and the dismantling of social structures.

Moreover, as *The Leadership Challenge* by Kouzes and Posner (2017) highlights, true

leadership in times of crisis requires a focus on the well-being of individuals and the collective organization. However, when war causes widespread destruction, leadership is forced to focus on mere survival rather than long-term flourishing, undermining any possibility of rebuilding the business or the society in the aftermath.

The emotional toll is also significant. Employees and citizens who have experienced trauma may struggle with mental health issues, reducing productivity and harming the morale of a nation or company. In modern management theory, leaders are increasingly recognizing the importance of emotional intelligence and empathy, especially after a crisis. In such situations, leaders who fail to acknowledge the emotional impact of war or disaster risk alienating their people, preventing the society or business from recovering.

The Dangers of Bias and Unwise Counsel in Leadership

Dhritarashtra's downfall can largely be attributed to his susceptibility to the manipulative counsel of his brother-in-law, Shakuni, whose personal vendettas and thirst for power led to the kingdom's ruin. In the *Sabhaparvan* of the *Mahabharata*, Shakuni orchestrates the infamous game of dice, which results in the Pandavas' exile, setting in motion the series of events that ultimately lead to the Kurukshetra War. His manipulations exploit Dhritarashtra's emotional attachment to his sons and his desire to maintain familial power, blinding him to the wise counsel of advisors like Vidura and Bhishma. Despite their warnings, Dhritarashtra approves the dice game, a decision that precipitates the war and the destruction of his kingdom.

This manipulation highlights a critical flaw in Dhritarashtra's leadership: his failure to seek impartial and wise counsel, choosing instead to trust a deceitful advisor who played on his biases and personal desires. Krishna's teaching in the *Bhagavad Gita* underscores the importance of leadership grounded in objective understanding and the ability to rise above personal attachments and biases: "यसनेषुमनुयाणां बुधि नाशो भवत्यु । तमासर्वं यनेन कार्यं बुधि परि यम् ॥ " (*Yasaneṣu manuṣyāṇāṃ buddhināśo bhavatyuta, Tasmāt sarvaprayatnena kāryaṃ buddhi-parikṣayam.*)—*Mahabharata, Udyogaparvan 43.36*. ("In times of crisis, a person's wisdom is often lost. Therefore, one must make every effort to preserve their intelligence and judgment.") Dhritarashtra's tragic fate serves as a warning against allowing emotions to cloud judgment, a lesson that remains relevant in both political and business leadership today. Shakuni's manipulative tactics mirror the concept of groupthink, as discussed by Irving Janis (1972), where leaders surrounded by sycophantic advisors make poor decisions due to their failure to encourage critical dissent. In Dhritarashtra's case, his reliance on Shakuni and his dismissal of wise counsel led to disastrous outcomes, similar to how modern leaders, influenced by advisors who reinforce their biases, can make flawed decisions.

The groupthink phenomenon, where dissenting opinions are suppressed in favor of conformity, can lead to the downfall of entire organizations or political systems. A contemporary example of this can be seen in the downfall of FTX in 2022, where Sam Bankman-Fried surrounded himself with close associates who failed to challenge his unsustainable financial strategies, leading to one of the largest cryptocurrency collapses in history. Similarly, in politics,

the mismanagement of the UK government under Prime Minister Liz Truss in 2022, where her economic policies led to financial turmoil, highlights the dangers of ignoring diverse viewpoints and objective counsel.

Furthermore, the concept of transformational leadership, emphasized by Bass (1990), stresses the importance of diverse perspectives and open communication. Dhritarashtra's failure to embrace these leadership qualities—by listening to Shakuni and ignoring other advisors—directly contributed to the downfall of Hastinapur.

From an economic standpoint, businesses that fail to encourage constructive criticism often suffer severe consequences. The collapse of WeWork in 2019, largely due to former CEO Adam Neumann's unchecked decision-making and disregard for investor concerns, illustrates how poor leadership and ignoring sound financial advice can lead to a company's downfall. The *Mahabharata* and modern management theories both highlight the critical need for leaders to seek impartial, objective counsel, encourage diverse viewpoints, and make decisions that serve the greater good. When leaders fail to do so, they risk not only their own downfall but also the destruction of the institutions they govern.

Dhritarashtra's Knowledge of the Opposition and His Failure to Act

Despite being aware that the Pandavas were righteous, powerful, and had Lord Krishna as their charioteer, Dhritarashtra allowed the war to unfold. In the *Sabhaparvan* of the *Mahabharata*, he was repeatedly warned by wise figures like Vidura and Bhishma about the catastrophic consequences of the impending conflict. Even Lord Krishna, who was aligned with the Pandavas, spoke to Dhritarashtra with profound clarity, urging him to abandon his attachment to his sons and consider the greater good. Yet, despite all these warnings and the moral clarity of the situation, Dhritarashtra's decision was swayed not by reason or righteousness but by his personal desires and biases, particularly his unwavering loyalty to his son Duryodhana. In one critical moment, he even attempts to stop the war but does so half-heartedly, driven by guilt, fear, and attachment rather than genuine concern for the kingdom's welfare. This failure to see beyond his personal attachments and recognize the moral righteousness of the Pandavas underscores a major flaw in his leadership: his inability to detach from personal bias and make objective decisions for the greater good. The *Mahabharata* highlights that leadership founded on such emotional attachments, rather than reason, leads to tragic consequences. Sri Aurobindo echoes a similar sentiment in his work *Essays on the Gita*, where he states: "*A leader must not be driven by personal affections and biases but by a higher vision that embraces the welfare of all. True action comes from a consciousness that is free from ego and attuned to the divine will.*" In leadership, whether in war or business, it is crucial to conduct a thorough analysis of the strengths, weaknesses, opportunities, and threats (SWOT) before making any significant decision. "बुद्ध्या युक्तो यः पश्यति शक्तिं स्वपरस्य च। स सफलं कर्म कुरुते न च मोहाय कल्पते॥" (*Buddhyā yukto yaḥ paśyati śaktim svaparasya ca, Sa saphalaṁ karma kurute na ca mohāya kalpate.*)—*Bhagavad Gita* 18.30 ("One who, with wisdom, assesses both his own strengths and those of his opponents performs his duties successfully and is not deluded by

confusion.")

A wise leader must assess their strengths, weaknesses, opportunities, and threats to make informed and strategic choices. Dhritarashtra's decision to go to war, despite knowing the likely defeat and the resulting devastation, starkly illustrates the consequences of failing to perform such an analysis. While he was aware that the Pandavas had righteousness, power, and Lord Krishna on their side, Dhritarashtra chose to ignore these factors and proceed with a path driven by personal attachment and bias. This highlights the failure to consider both the internal (strengths and weaknesses) and external (opportunities and threats) elements that would have helped him understand the broader consequences of war.

This failure can be seen in modern history. The 2022 Russia-Ukraine conflict, where strategic miscalculations led to prolonged warfare and economic sanctions, mirrors Dhritarashtra's misjudgment. Similarly, in business, the downfall of Evergrande in 2023 due to reckless financial decisions highlights the cost of ignoring financial stability and risk assessment. The collapse of Silicon Valley Bank in 2023 serves as another example of leadership negligence, where failure to anticipate market risks and economic instability led to a banking crisis.

Importance of Righteousness in Leadership

The concept of righteousness is central to the teachings of the Mahabharata and the Bhagavad Gita, where it governs both individual conduct and the principles of leadership. Leaders are expected to act according to ethical principles that promote justice, fairness, and the welfare of their people, ensuring that their decisions serve the common good rather than personal interests. In the Mahabharata, King Dhritarashtra's failure to uphold righteousness is a key factor in the ruination of his family and kingdom. His blind loyalty to his sons, especially Duryodhana, led him to disregard the greater good of his people, resulting in the catastrophic Kurukshetra War, the destruction of Hastinapur, and the suffering of countless individuals. This failure exemplifies the importance of leaders acting in accordance with dharma—something that is often overlooked when personal attachments or biases cloud their judgment.

The Bhagavad Gita provides numerous teachings on leadership and duty. One of the key lessons is performing one's duty without attachment to the results. As stated in Bhagavad Gita 3.19: “तस्मादसक्तः सततं कार्यं कर्म समाचर | असक्तो ह्याचरन्कर्म परमाप्नोति पूरुषः ||” (Therefore, always perform your duty without attachment. By working without attachment, one attains the Supreme). Ethical business leaders, such as Ratan Tata, exemplify this by prioritizing the welfare of society over short-term profits. His commitment to philanthropy and ethical business practices reflects the ideal of working for the greater good without selfish attachment.

In the realm of business, politics, and economics, ethical leadership determines long-term success. Companies that prioritize dharma-like ethical principles, such as Infosys under Narayana Murthy, demonstrate that integrity and transparency build sustainable organizations. Similarly, economic policies should serve the collective good. The Scandinavian model of social democracy, which balances capitalism with welfare economics, aligns with the dharmic principle of ensuring the welfare of all rather than favoring a select few. The Bhagavad Gita also teaches

that war must only be fought for justice and the protection of righteousness. Leaders must consider whether conflicts are just and ethical. For instance, Winston Churchill's resistance against Nazi Germany was based on protecting the larger good against tyranny.

To ensure that leaders embody dharma, education in ethics and righteousness must be integral to leadership development. Understanding ethical decision-making, fairness, and justice should be a fundamental part of leadership training, whether in political governance, corporate management, or policymaking. Leadership programs should integrate teachings from scriptures like the Bhagavad Gita alongside modern leadership theories to cultivate wise, just, and ethical leaders. In the Mahabharata, Yudhishtira, known for his commitment to dharma, constantly sought guidance from wise advisors like Vidura and Bhishma. Similarly, modern leaders should undergo systematic learning and mentorship in dharma, ethics, and governance to ensure that they lead with wisdom and righteousness. Exposure to such teachings will help leaders navigate complex moral dilemmas and make decisions that benefit society as a whole. Leaders who follow dharma in both their personal and professional lives are more likely to create stable, prosperous, and just environments.

By embracing dharma, leaders in all fields can ensure their decisions lead to the prosperity, stability, and well-being of society as a whole.

Yudhishtira's Restoration of the Kingdom's Glory

After the devastation of the Kurukshetra War, Yudhishtira, the eldest Pandava, ascended to the throne of Hastinapur. The kingdom lay in ruins: many of its citizens were dead, the economy had collapsed, and despair hung in the air. Yet, Yudhishtira's leadership brought a new sense of hope to the devastated land. Guided by the principles of righteousness and by the wise counsel of figures like Bhishma and Vidura, Yudhishtira worked tirelessly to restore the kingdom. He prioritized the welfare of his people and rebuilt the economy, demonstrating that ethical leadership can heal even the deepest wounds of a society.

Yudhishtira's commitment to justice, fairness, and selfless service exemplified the transformative power of righteous leadership. His governance shows that in the aftermath of destruction, he is the ethical leader who can inspire a nation toward recovery and growth, rebuilding not only the material wealth but also the moral fabric of society. This is reflected in the Bhagavad Gita 2.21: “वेदाविनाशिनं नित्यं य एनमजमव्ययम् | कथं स पुरुषः पार्थ कं घातयति हन्ति कम् ||” (He who knows the self to be indestructible, eternal, unborn, and unchangeable—how can that person slay or cause another to slay?). Yudhishtira's resilience and adherence to dharma despite immense suffering reflect the principle that righteous leaders act with a deep sense of duty and responsibility, rather than being swayed by transient emotions or personal losses.

In modern leadership theory, Yudhishtira's leadership can be understood through the lens of transformational leadership, a concept articulated by James Bass (1990). Bass defines transformational leadership as a style in which leaders inspire followers by creating a vision of the future and leading through personal integrity, selflessness, and commitment to the common

good. Yudhishtira's actions after the war reflect these principles, as he worked to restore the kingdom not for personal gain, but for the collective well-being of his people. Furthermore, his leadership embodies the essence of ethical decision-making, which is a key theme in contemporary management literature. Kouzes and Posner (2017), in *The Leadership Challenge*, argue that leaders who prioritize the values of integrity, responsibility, and accountability can achieve not only business success but also positive social impact. Yudhishtira's story provides a timeless reminder that no matter how dire the circumstances, ethical leadership grounded in dharma can lead to recovery and lasting growth.

In the modern world, ethical governance is essential for long-term prosperity, as demonstrated by leaders across various fields. In politics, figures like Nelson Mandela exemplified the power of ethical leadership by prioritizing reconciliation and justice over revenge, fostering national unity and healing. In business, Satya Nadella's leadership at Microsoft highlights the significance of empathy and innovation in driving corporate success while maintaining ethical responsibility. Similarly, in economics, sustainable development policies that focus on environmental responsibility and equitable wealth distribution align with the principles of dharma, ensuring long-term stability. A remarkable example of ethical business leadership is Howard Schultz, the former CEO of Starbucks, who revitalized the company during its decline in the late 2000s by prioritizing employee welfare, ethical sourcing, and customer experience over mere profit maximization. His commitment to corporate social responsibility helped Starbucks regain its reputation and financial strength. Likewise, Lee Iacocca, the former CEO of Chrysler, rescued the company from bankruptcy in the late 1970s by securing a government loan and restructuring it with a focus on quality, innovation, and customer needs, showcasing the impact of responsible corporate management. Just as Yudhishtira's adherence to dharma helped restore Hastinapur after the Kurukshetra War, modern leaders must embody these values to create a just and prosperous society. By embracing dharma, leaders across all domains can ensure that their decisions lead to the prosperity, stability, and well-being of society as a whole.

Conclusion

The narratives of King Dhritarashtra and Yudhishtira from the *Mahabharata* offer valuable lessons in leadership ethics. Dhritarashtra's biased decisions, influenced by personal attachments, led to the downfall of Hastinapur, emphasizing the dangers of unethical leadership. On the other hand, Yudhishtira's commitment to dharma, along with his consultative approach to governance, highlights the importance of moral leadership in restoring peace and stability. His leadership exemplifies transformational qualities that prioritize the welfare of the people. As the Bhagavad Gita (3.21) states:

"Yad yad ācarati śreṣṭhas tat tad evetaro janāḥ | sa yat pramāṇaṁ kurute lokas tad anuvartate ||"
(Whatever action a great person performs, common people follow. Whatever standards they set, the world pursues.)

This wisdom resonates with modern leadership theories, which emphasize the need for ethical decision-making and the well-being of followers. Bass (1990) asserts that transformational leaders inspire change by maintaining integrity and empowering others. Kouzes and Posner (2012) emphasize that exemplary leaders "model the way" by aligning their actions with shared values, which fosters trust and commitment within organizations. These modern concepts echo the lessons of the *Mahabharata*, where Yudhishtira, despite his strong sense of dharma and wisdom, seeks guidance from figures like Bhishma and Vidura on how to rule ethically and justly. Bhishma's teachings in the *Shanti Parva* emphasize the responsibility of a ruler to prioritize the welfare of their subjects over personal interests, a principle that resonates in today's leadership practices.

In contemporary examples of ethical leadership, Jacinda Ardern, the former Prime Minister of New Zealand, is a notable figure. Her compassionate and transparent leadership during crises, such as the Christchurch attacks and the COVID-19 pandemic, helped reinforce public trust and stability. Her approach reflects the ancient Indian ideal of *Sevāvṛtti* (selfless service), showing how ethics and empathy in leadership can create positive societal outcomes. In the business world, Patagonia's founder, Yvon Chouinard, provides another example of ethical leadership. He built a company based on environmental responsibility and corporate social responsibility, emphasizing sustainability and employee well-being over short-term profits. Patagonia's success illustrates that moral leadership in business not only benefits society but also fosters long-term success and brand loyalty.

Many countries have recognized the importance of ethical leadership and have implemented training programs to develop such leadership in politics, governance, and business. For example, Singapore's Civil Service College offers structured leadership training for government officials, focusing on ethical governance and policy making. Similarly, Scandinavian countries like Norway and Sweden incorporate ethical leadership and transparency into their public administration training. Harvard Business School and INSEAD also provide executive programs to equip leaders with skills for responsible decision-making, ensuring they contribute to sustainable global practices.

On the flip side, unethical leadership practices have resulted in corporate collapses, such as the Enron scandal, where deceit and manipulation led to financial ruin. This mirrors the failure of Dhritarashtra, who, by failing to restrain his son, precipitated catastrophic consequences for Hastinapur.

From a political perspective, the leadership of figures like Nelson Mandela exemplifies Yudhishtira's wisdom. Mandela's reconciliation and commitment to ethical governance played a crucial role in healing a divided nation. Conversely, contemporary conflicts, such as the Russia-Ukraine war, illustrate the dire consequences of leadership driven by personal ambition rather than collective welfare, further reinforcing the lessons of the *Mahabharata*.

However, it is important to recognize that the *Mahabharata* is not a direct historical account but a story told through the lens of the victors, particularly from the perspective of the

Kshatriya dynasties. The narrative has been shaped by multiple layers of storytelling, including its telling to Janamejaya, the great-grandson of Arjuna, by a bard at Shaunaka's Ashrama. This retrospective framing suggests that the events we receive today have been filtered through various perspectives, each adding and subtracting from the original accounts. Despite the fact that the *Mahabharata* may not provide a precise historical record, we are grateful for the lessons it imparts. The story teaches us the importance of righteousness and the inevitable downfall of unrighteousness. Even though intricate details may have been lost in the retelling, the essence of the narrative continues to shape our values, offering guidance on how to live ethically and lead with integrity.

As the *Manusmriti* (7.92) states: "*Na rājā rājyam ādadyāt svārtham ullaṅghya dharmataḥ | prajāḥ pālayan rājā dharmam pālayate svayam ||*" (*A king should not exploit his kingdom for selfish gain but should uphold dharma by protecting his people.*)

The lessons of the *Mahabharata* remind us that leadership rooted in ethics, empathy, and a commitment to the collective well-being leads to sustainable success and societal harmony. As we look to modern examples in politics, business, and governance, it is clear that these principles continue to guide effective leadership today. By balancing personal integrity with a focus on the greater good, leaders can create environments where both individuals and societies thrive.

Works Cited

- Sri Aurobindo. (1920). *The Human Cycle and The Foundations of Indian Culture*. The Arya Press.
- Aurobindo, S. (1922). *Essays on the Gita*. Sri Aurobindo Ashram Publication Department.
- Janis, I. (1972). *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Houghton Mifflin.
- Van Buitenen, J. A. B. (Trans.). (1975). *Mahabharata: Volume 1, Sabhaparvan & Udyogaparvan*. University of Chicago Press.
- Christensen, C. M. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Review Press.
- Mahabharata, translated by C. Rajagopalachari (2000), Bharatiya Vidya Bhavan.
- Kautilya. (2000). *Arthashastra: The Science of Political Economy*. Translated by L. N. Sharma. Delhi: Motilal Banarsidass Publishers.
- Mugabe, R. (2001). *The Economic and Political Collapse of Zimbabwe: A Case Study in Poor Leadership*. Global Policy Journal.
- Greenleaf, R. K. (2002). *Servant Leadership: A Journey into the Nature of Legitimate Power and Greatness*. Paulist Press.
- Enron Corporation Collapse. (2003). *Case Study of Corporate Governance Failures*. Harvard Business Review.
- Carmignani, F. (2003). *The Economics of War: Impacts on Trade, Infrastructure, and National Growth*. World Economics Review.
- Fitzgerald, J. (Trans.). (2004). *Mahabharata: Striparvan*. University of Chicago Press.
- Sivananda, S. (2004). *The Mahabharata: The Essence of the Epic*. Divine Life Society.
- Bhagavad Gita, translated by Eknath Easwaran (2007), Nilgiri Press.
- Dixit, A. K., & Nalebuff, B. J. (2008). *The Art of Strategy: A Game Theorist's Guide to Success in Business and Life*. W.W. Norton & Company.
- Achenbaum, W. A. (2009). *Leadership and the Economy: The Impact of Personal Bias in Leadership*. Economic Leadership Journal.
- Hudson, E. (2010). *Disorienting Dharma: Ethics and Suffering in the Mahabharata*. Oxford University Press.
- UNESCO. (2011). *Long Walk of Peace: An Examination of Peace and Conflict Resolution Practices*. UNESCO Publications.
- Kouzes, J. M., & Posner, B. Z. (2012). *The Leadership Challenge* (5th ed.). Jossey-Bass.
- Hudson, E. (2013). *Disorienting Dharma: Ethics and the Mahabharata*. Oxford University Press.
- Bishop, S. (2015). *Ethical Leadership in the 21st Century*. Harvard Business Review.
- Kouzes, J. M., & Posner, B. Z. (2017). *The Leadership Challenge: How to Make Extraordinary Things Happen in Organizations*. John Wiley & Sons.

- Goodwin, D. K. (2018). *Leadership in Turbulent Times*. Simon & Schuster.
- Schilling, M. A. (2020). *Strategic Management: Theory and Cases*. Pearson Education.
- Manusmriti (7.92), translated by Patrick Olivelle, Oxford University Press.

Long-Term Effects of Thymectomies and Congenital Heart Disease on Immune Function and T-Cell Composition By Aarush Anand

Abstract

Congenital Heart Disease (CHD) has affected humanity for centuries, and only recently have we discovered treatments for this condition. However, one of the treatments that is very commonly used in CHD patients requires a thymectomy (or removal of the thymus). This research paper investigates how CHD influences T-cell composition and function after infant thymectomies, focusing on immune-related complications in affected patients as well as the impact of non-thymectomized CHD on immune cell compositions. An extensive literature review of 19 scientific papers, including articles from PubMed, Science Direct, and NIH, was conducted. These studies, published between 2005 and 2024, highlight the specific T-cell subtypes impacted by CHD and the mechanisms driving related health complications. This analysis revealed that full or partial thymectomies performed for CHD-related complications resulted in decreased or changed subsets in Natural Killer (NK) cells, Cytotoxic T-cells (CD8+), Helper T-cells (CD4+), and other prevalent immune cells. Studies also found that CHD patients have an increased risk of several complications, including SARS-CoV-2 (COVID-19), Influenza, and Bronchopneumonia, regardless of whether they had a thymectomy or not. However, early vaccination efforts have proven to improve the prognosis for CHD patients for diseases like influenza and can reduce the risk of adverse events from such complications. Future thymectomies for CHD patients should include such preventative measures to reduce the risk of life-threatening complications from happening because of the lack of T-cells.

Introduction

Congenital Heart Disease (CHD) is the most common birth defect in the world, affecting over 13 million people. Although advances in surgical interventions for CHD have improved survival rates, their long-term effects on immune health are often underexplored. Infants who have structural defects due to CHD have to undergo surgery to ensure oxygen-rich and oxygen-poor blood stay in their respective chambers. While these procedures greatly help the infant survive, they commonly involve a partial or complete thymectomy (a removal of the thymus). Thymus removal is preferred by pediatric cardiothoracic surgeons since it allows access to key areas of the heart (such as the aorta, in pulmonary arteries, and superior/inferior vena cava) and improves visibility, giving better surgery results when operating on the structural defects. Unfortunately, this also hinders immune system health, as the thymus is one of the most important immune organs for the maturation and differentiation of new T-cells, including naïve, CD4+, CD8+, and many more. Since they have matured most of their T-cells, Adults with CHD (AwCHD) patients aren't affected by thymectomies as much and rarely suffer future immune

dysfunctions because of it. However, in infants and newborns, it can cause long-term immune dysfunction and increased susceptibility to adverse complications, as their thymus hasn't trained enough T-cells to provide sufficient long-term immunity against pathogens. Aside from thymectomy, CHD can still prove to be detrimental to immune system health. Many studies have found that CHD can cause a decrease in certain immune cell subsets, which predisposes them to many diseases such as Bronchopneumonia.

This review focuses on CHD and the role of infant thymectomy on T-cell composition and long-term immune function. This area is often overlooked when treating patients with CHD, and this review aims to shed more light on the future consequences of thymectomies in patients with CHD to long-term immune health to improve patient prognosis. This paper also aims to provide more clarity about the different complications that may arise due to CHD with or without thymectomy to help pave the way for further treatment options for patients with CHD.

Methods

A comprehensive investigation through reliable platforms such as Google Scholar and PubMed was conducted. The 19 studies included in this review were published on or after 2005. Literature reviews and case studies were utilized to synthesize existing evidence and recent discoveries. The study includes scientific articles relating to T-cell compositions and risk factors for AwCHD. Keywords such as “Congenital Heart Disease,” “T-cell Composition,” “thymectomy,” and “immune” were used to gather relevant information. Scientific articles were included only from reputable, peer-reviewed journals, such as *NIH*, *ScienceDirect*, *Cambridge Core*, *Research Gate*, *Nature*, *Springer Nature*, *Healthline*, *Journal of Allergy and Clinical Immunology*, and *The Journal of Clinical Investigation*. All other sources have been omitted from this review.

From the articles referenced, the abstract, discussion, and results were used, taking into account mentions of T-cell composition or immune complications. This was done to ensure the proper parsing of the text and the drawing of accurate conclusions. The different types of T-cells affected were also noted. To organize this data, a document with summaries of relevant information from each publication was compiled. It included key terms and ideas that were useful when compiling the results. Only Natural Killer (NK), CD8+, and CD4+ cells were analyzed and included in this article; similarly, COVID-19, influenza, and Bronchopneumonia were analyzed for complications at a higher risk. This does, however, leave out key details regarding other possible pathogens and T-lymphocyte subsets impacted by CHD and treatment options like thymectomy.

Results

CHD has been shown to impact various subsets of immune system cells. The analysis of the articles cited shows that CHD affects NK, CD8+, CD4+, and other subsets of immune cells.

Natural Killer (NK) Cells

NK cells play a profound role in the immune system and are usually referred to as the bridge between innate and adaptive immunity. NK cells are responsible for destroying infected and diseased cells, including cancer cells. They also interact with adaptive immune cells, such as macrophages and dendritic cells, allowing for stronger immune defense (Vivier et al., 2008). Deficiency and inactivity of these cells can cause an increased risk of cancer and other microbial diseases.

There are 2 prominent types of NK cells in the body: CD56^{dim} and CD56^{bright} NK cells. CD56^{dim} NK cells are known to be more cytotoxic and have low cytokine production. They make up around 90% of circulating NK cells in the human body and are responsible for killing infected cells. CD56^{bright} NK cells on the other hand make up around 10% of NK cells and are responsible for the production of cytokines, specifically interferons (Moretta, 2010). One study found that CD56^{bright} were elevated while CD56^{dim} cells stayed the same in AwCHD, typically more prevalent in further stages of CHD (Wienecke, Cohen et al.). CD56^{bright} NK cells are known to be inflammatory cells, releasing cytokines that cause inflammation. This suggests that thymectomy in CHD patients could lead to chronic inflammation by disrupting the homeostasis of NK cells, specifically due to the increase in CD56^{bright} cells. Since CD56^{bright} NK cells release significant amounts of cytokines, their increase in population may contribute to inflammatory responses, regardless of age.

Although NK cells aren't trained or matured by the thymus (since they are part of the innate immune system), they are still affected by infant thymectomy. One study found that after pediatric patients underwent thymectomies during cardiac surgery, NK populations had significantly changed, suggesting that the thymus indirectly influences the maturation of NK cells, although this relationship needs to be studied further. This study also showed that infant thymectomy, in comparison to thymectomy at an older age, had a lack of normal NK differentiation (Halnon et al., 2005). Another study found that NK receptor expression was upregulated in CHD patients with thymectomies, causing an increase in inflammation and tissue damage due to the constant activity of the NK cells (Cramer et al., 2024). Although this study didn't state which types of NK cells caused the inflammation, it is likely that the CD56^{bright} cells caused it due to their inflammatory properties and cytokine production. The increased activity and cytokine production of the CD56^{bright} cells likely caused the inflammation in the patients.

Cytotoxic T-cells (CD8+)

CD8+ T-cells, otherwise known as cytotoxic T-cells, are responsible for the identification of MHC Class 1 molecules using the CD8 receptor to recognize infected cells. The cells then release cytokines or small proteins that ignite an immune response, to fight the infection. A lack of these cells has been proven to cause severe immunodeficiency, and autoimmune and autoinflammatory disorders (Liszewski et al., 2016).

One source found that CHD thymectomy impaired production, and resulted in long-term decrease in CD8+ lymphocytes. This article also demonstrated the difference in lymphocyte counts between partial and full thymectomy. Specifically, it found that partial thymectomy led to

higher counts of T-lymphocyte populations compared to a full thymectomy, as shown in the graph below (Halnon et al., 2005).

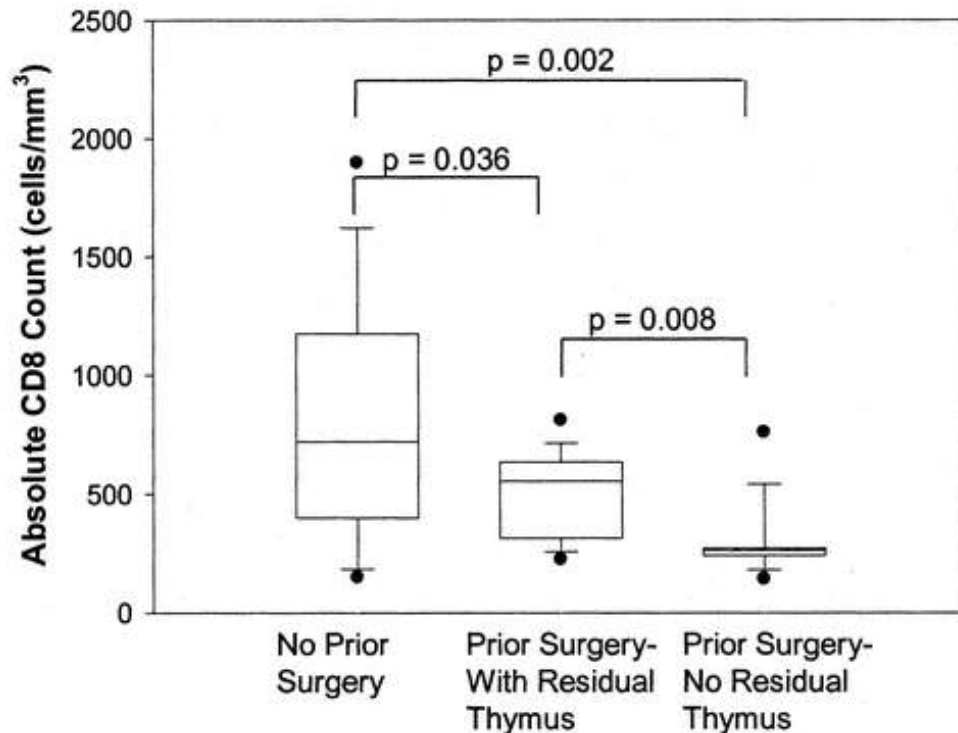


Fig. 1. Frequencies of CD8+ lymphocytes compared to surgical status (Halnon et al., 2005).

From the results shown in the graph, it can be concluded that partial thymectomy would help preserve immune health and lymphocyte count much more compared to a full thymectomy, something that pediatric cardiothoracic surgeons should keep in mind when performing surgery on a patient. This could help decrease the chances of several complications from occurring in the future, such as Bronchopneumonia which is, as discussed later in this article, at a higher risk of developing in CHD patients.

Additionally, it was found that CHD in general, regardless of thymectomy, caused lower amounts of circulating CD8+ cells. However, this count could also be influenced by heart failure in general, that is caused by CHD (Wieneke, Lemke et al., 2022). Indeed, this count may have been influenced by thymectomy, however its effect wasn't measured as part of the study. Regardless, this study helped clarify how CHD impacts the immune T-lymphocyte subsets found in children and adults, providing additional evidence on the negative effects of CHD on immune health. In addition to this, another article found that neonatal thymectomy reduced CD8+, which is not seen in thymectomy conducted later in life (Broek et al., 2016). This further confirms the effects of thymectomy on immune health in infants, as they haven't matured/created the majority of their T-lymphocytes, causing significant reductions compared to healthy controls.

Helper T-cells (CD4+)

Another very important type of cell, crucial to the optimal function of the immune system, is helper T-cells, or CD4+ cells. There are many categories of helper T-cells, however, the general purpose of CD4+ cells is to activate and coordinate other immune cells during an immune response. A lack of these cells causes a delayed immune response, causing severe tissue damage among other consequences (Luckheeram et al., 2012).

Similar to CD8+ cells, CD4+ cells experience a drop in total circulating population in CHD patients, with more cells being lost with a full thymectomy compared to a partial thymectomy. This study also found that the age at which the individual received a thymectomy did not affect the lymphocyte count (Halnon et al., 2005). Additionally, it was found that low CD4+ lymphocytes led to increased risk of adverse cardiovascular events (Wieneke, Lemke et al., 2022).

One source found that the CD4+ count initially decreased following a thymectomy. However, after the regrowth of the thymus, the naive CD4+ count increased, allowing for significantly improved population count (Broek et al., 2016). This evidence points to the possibility of thymus regrowth to positively impact patient prognosis.

COVID-19

COVID-19 emerged during the 2020 pandemic, and thanks to vaccines and other therapies, we have been able to successfully treat it. However, for many CHD patients, this disease could cause significant health risks. In general, CHD increases the risk of negative outcomes when the patient is infected (Haiduc et al., 2020).

According to one study, AwCHD with T-cell senescence and chronic inflammation tend to have higher risk of serious complications after a COVID-19 infection, although it is mainly the elderly population that are impacted. In addition, patients with Down Syndrome and CHD tend to have a much more severe form of the disease after a COVID-19 infection (Wieneke, Cohen et al., 2021). These studies show that COVID-19 can cause negative complications in CHD patients, which is further exaggerated by the removal of the thymus. Especially since COVID-19 is very contagious, the severe lack of T-cells caused by the thymectomy leads to an underdeveloped immune response, with CD4+, CD8+, and CD56+ (or NK cells) having a much more difficult time creating a sufficient immune response.

A case-study reviewing a 6-month-old child with CHD and who was infected with COVID-19 found that the child had an increased amount of activated T-cells and lower amounts of naive T-cells. However, this count was skewed towards memory T-cells, which showed the functioning of the adaptive immune system (Wurzel et al., 2021). This study was done without a thymectomy of the child, so T-lymphocyte populations weren't greatly influenced and showed that CHD patients still have some immunity if they retain their thymus.

Influenza

The seasonal influenza consistently changes each year, with different variants occurring due to the very rapid mutation rate of the virus. The pathogen is already very dangerous, especially in children, and with CHD, the risk is increased significantly. Many AwCHD have

higher risks of cardiovascular complications once exposed to influenza, and it is recommended that cardiologists suggest influenza vaccines to AwCHD patients (Yeh et al., 2024).

Although influenza can be very deadly in CHD patients, especially due to the lack of key T-lymphocytes, vaccination efforts have shown to be successful in providing a normal T-cell response when CHD patients are infected with the influenza virus (Leone et al., 2022). However, this is only when the patient is in young adulthood, not childhood or elder years, which still provides a dangerous risk for those age groups. Additionally, another study showed that the dosage at which the influenza vaccine is given doesn't impact the effectiveness of the vaccine, showing that even small amounts of the vaccine could be useful in helping provide preventative measures against harmful diseases, shown in the graph below (Dehghani et al., 2024).

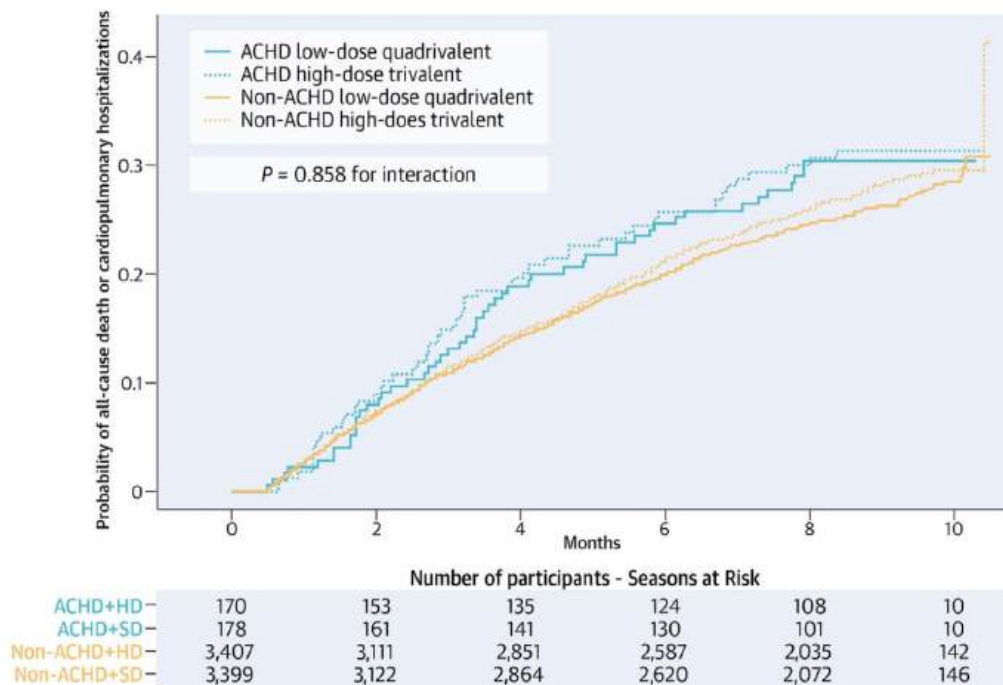


Fig. 2. Probabilities of adverse cardiovascular events after Influenza virus in CHD patients (Dehghani et al.).

The influenza vaccine comes in many different forms, some being attenuated (Wong et al., 2013). Lower dosages can be useful for reducing the risk of an infection from occurring from the live virus, especially important given the already immunocompromised state the patient with CHD lives in. With this information, we can continue testing on other infectious pathogens and provide preventative measures for many viral pathogens for CHD patients in the future. Nevertheless, further lab testing and clinical trials are needed for other viruses to confirm the effectiveness of the vaccines on AwCHD.

Bronchopneumonia

Bronchopneumonia is a bacterial infection affecting the bronchioles and alveoli of the lungs, causing severe swelling. Bronchopneumonia is known to cause fever, chest pain, and

headaches. However, in children, this can be severe, and can cause a significant drop in blood pressure, rapid heart rate, and retractions of chest muscles (Martel, 2019).

Evidence has shown that bronchopneumonia is more common in CHD patients, especially children. One experiment that compared the differences in T-lymphocyte composition in healthy controls compared to children with CHD found that the CHD patients had lower levels of CD4+, CD8+, and many other subsets of lymphocytes (Huang et al., 2014).

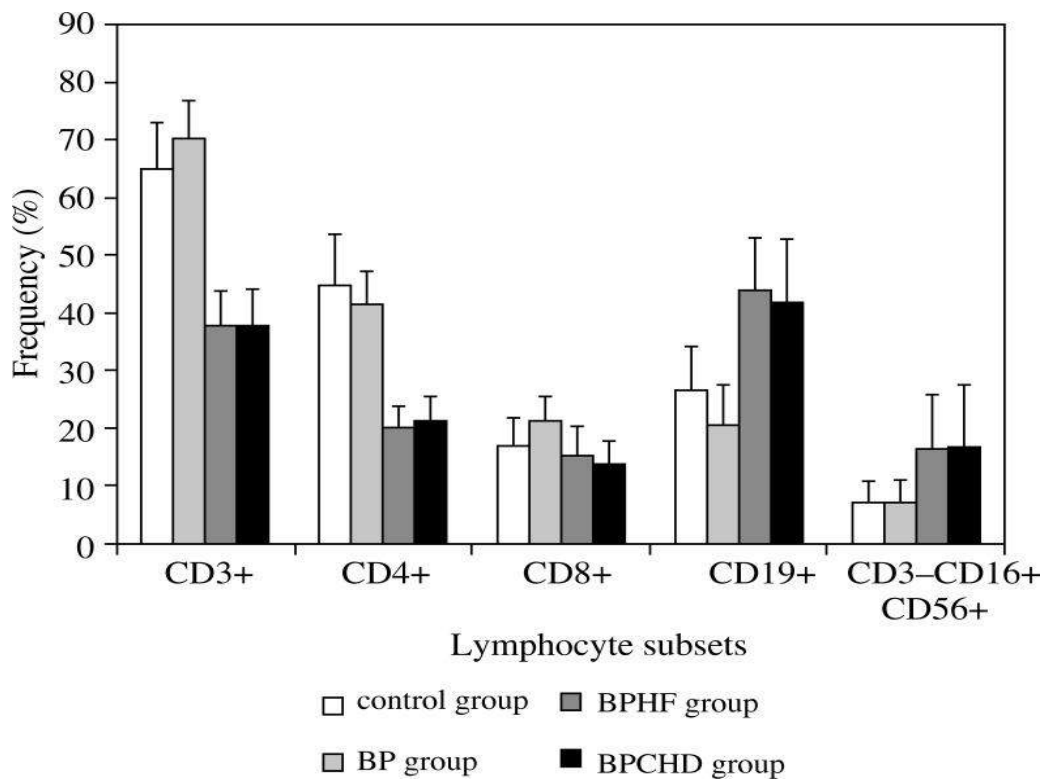


Fig. 3. The frequencies of Lymphocyte subsets in each group of subjects (Huang et al., 2014).

As shown in the graph, many key T-lymphocytes were decreased in bronchopneumonia patients with CHD, including CD3+, CD4+, and CD8+. The increase in CD19+, CD16+, and CD56+ may be a compensatory immune response due to both CHD-related immune alterations and the body's reaction to bronchopneumonia.

Additionally, it was found that CHD patients before thymectomy, had immune complications even with the thymus. The study found that there was impaired T-cell development and thymic atrophy in patients with CHD (Bremer et al., 2024) This study shows that it is imperative that we take into account the dangers of CHD on the immune system, as thymus problems could further severe complications in the future.

Discussion

These results show the need for more awareness regarding the lack of immunity in CHD patients. Future surgeries for structural CHD defects should aim to do a partial thymectomy

instead of a full one to preserve immune health. However, as discussed earlier, even partial thymectomy can lead to altered immune compositions and therefore cause immune dysfunction, but a lot less compared to a full thymectomy (Halnon et al., 2005). Currently, alternative methods for surgically fixing the defects found in many CHD patients haven't been found, making it difficult for a drastic change in these procedures. This is because of the location of the heart, since many other angles for accessing the heart are too dangerous for surgical intervention.

One way to preserve immune function in CHD patients would be to ensure the proper vaccine administration for many of these infectious diseases to help familiarize the immune system with these diseases and increase quality of life. However, this method is underexplored and hasn't been proven yet with adequate scientific evidence, thus calling for more clinical trials to prove this method's efficacy.

Another possible option that may be available for thymectomized CHD patients is the regrowth of the thymus. Early clinical studies have shown regrowth of the thymus to its original, peak function through a process called thymic involution (Chaudhry et al., 2016). However, this process is still being developed and likely won't be available for many years to come. Even when it does become available for the general public, it is likely to be very expensive and cause a socioeconomic difference in treatment options. Nevertheless, this would offer a great solution for these patients and also provide the best possible surgery results for the individuals in the future.

Although this review went over multiple T-cell compositions, it fails to address the many other immune cells affected. This does call for more research to be done on the compositions of other cells such as eosinophils and basophils to provide a more detailed prognosis for those who suffer from CHD. Additionally, this review didn't include other diseases that are at an increased risk in developing in CHD patients. Since long-term prognosis predictions are limited, having increased knowledge of the possible diseases that can cause negative complications would allow for the better preparation and treatment plan for thymectomized CHD patients. In light of this, more studies should be performed on other possible diseases at higher risks, specifically longitudinal studies on thymectomized patients.

In summary, this literature review provides valuable insight into the factors at play in CHD patients that cause a lack of immunity. However, further research into this underexplored field is needed, as many gaps in knowledge still exist. Addressing these gaps will allow for the development of novel therapies and treatments to help combat these dangerous, and sometimes fatal, complications for thymectomized and non-thymectomized patients.

Conclusion

This review underscores the critical role of immune health in the management of both thymectomized and non-thymectomized CHD patients. Specifically, this review emphasizes the impact of CHD on T-cell compositions, finding that CHD patients lowered/altered compositions of NK, CD8+, and CD4+ cells as well as increasing the risk of several complications including COVID-19, Influenza, and Bronchopneumonia, due to the weakened immunity from the lowered

T-cell subsets. With these findings, physicians treating CHD patients should take into account the impaired immunity when treating them to ensure the most optimal care.

Works Cited

- Bremer, Sarah-Jolan, et al. "Thymic Atrophy and Immune Dysregulation in Infants with Complex Congenital Heart Disease - Journal of Clinical Immunology." SpringerLink, Springer US, 23 Feb. 2024, doi.org/10.1007/s10875-024-01662-4.
- Broek, Theo van den, et al. "Neonatal Thymectomy Reveals Differentiation and Plasticity within Human Naive T Cells." *The Journal of Clinical Investigation*, American Society for Clinical Investigation, 1 Mar. 2016, doi.org/10.1172/JCI84997.
- Chaudhry, Mohammed S, et al. "Thymus: The next (Re)Generation." *Immunological Reviews*, U.S. National Library of Medicine, 1 May 2016, doi.org/10.1111/imr.12418.
- Cramer, Alexa, et al. "Early-Life Thymectomy Leads to an Increase of Granzyme-Producing $\Gamma\delta$ T Cells in Children with Congenital Heart Disease." *Nature News*, Nature Publishing Group, 13 Nov. 2024, doi.org/10.1038/s41467-024-51673-3.
- Dehghani, Payam, et al. "Feasibility and Findings of Including Self-Identified Adult Congenital Heart Disease Patients in the INVESTED Trial." *Journal of the American College of Cardiology*, JACC Journals, 3 Apr. 2024, doi.org/10.1016/j.jacadv.2024.100897.
- Haiduc, Ana Alina, et al. "Covid-19 and Congenital Heart Disease: An Insight of Pathophysiology and Associated Risks: Cardiology in the Young." Cambridge Core, Cambridge University Press, 11 Nov. 2020, doi.org/10.1017/S1047951120003741.
- Halnon, Nancy J, et al. "Thymic Function and Impaired Maintenance of Peripheral T Cell Populations in Children with Congenital Heart Disease and Surgical Thymectomy." *Nature News*, Nature Publishing Group, 1 Jan. 2005, doi.org/10.1203/01.PDR.0000147735.19342.DE.
- Huang, Rui, et al. "Cellular Immunity Profile in Children with Congenital Heart Disease and Bronchopneumonia: Evaluation of Lymphocyte Subsets and Regulatory T Cells." *Central-European Journal of Immunology*, U.S. National Library of Medicine, 15 Dec. 2014, doi.org/10.5114/ceji.2014.47734.
- Leone, David M., et al. "T Cell Response to Influenza Vaccination Remains Intact in Adults with Congenital Heart Disease Who Underwent Early Thymectomy." *International Journal of Cardiology Congenital Heart Disease*, Elsevier, 25 Mar. 2022, doi.org/10.1016/j.ijcchd.2022.100359.
- Liszewski, W, and R Gniadecki. "The Role of Cytokine Deficiencies and Cytokine Autoantibodies in Clinical Dermatology." *Journal of the European Academy of Dermatology and Venereology : JEADV*, U.S. National Library of Medicine, 30 Mar. 2016, doi.org/10.1111/jdv.13303.
- Luckheeram, Rishi Vishal, et al. "CD4+T Cells: Differentiation and Functions - Luckheeram - 2012 - Journal of Immunology Research - Wiley Online Library." *CD4+T Cells*:

- Differentiation and Functions*, Journal of Immunology Research, 2012, doi.org/10.1155/2012/925135.
- Martel, Janelle. "Bronchopneumonia." *Bronchopneumonia: Symptoms, Risk Factors, and Treatment*, Healthline Media, 19 Apr. 2019, www.healthline.com/health/bronchopneumonia.
- Moretta, Lorenzo. "Dissecting CD56 Dim Human NK Cells - Ash Publications." *Dissecting CD56dim Human NK Cells*, 2010, doi.org/10.1182/blood-2010-09-303057.
- Vivier, Eric, et al. "Functions of Natural Killer Cells." *Nature Immunology*, U.S. National Library of Medicine, 9 May 2008, doi.org/10.1038/ni1582.
- Wienecke, Laura M., et al. "Altered Compositions of Monocyte, T Lymphocyte and NK Cell Subsets in Heart Failure of Adult Congenital Heart Disease." *International Journal of Cardiology Congenital Heart Disease*, Elsevier, 5 Sept. 2022, doi.org/10.1016/j.ijcchd.2022.100418.
- Wienecke, Laura M., Sarah Cohen, et al. "Immunity and Inflammation: The Neglected Key Players in Congenital Heart Disease? - Heart Failure Reviews." *SpringerLink*, Springer US, 2 Dec. 2021, doi.org/10.1007/s10741-021-10187-6.
- Wong, Sook-San, and Richard J Webby. "Traditional and New Influenza Vaccines." *Clinical Microbiology Reviews*, U.S. National Library of Medicine, 26 July 2013, doi.org/10.1128/CMR.00097-12.
- Wurzel, Danielle, et al. "A Case Report Describing the Immune Response of an Infant with Congenital Heart Disease and Severe COVID-19." *Nature News*, Nature Publishing Group, 15 Nov. 2021, doi.org/10.1038/s43856-021-00047-7.
- Yeh, Doreen DeFaria, and Charmaine S. Lam. "Influenza Vaccination in Adult Patients With Congenital Heart Disease." *Journal of the American College of Cardiology, JACC Journals*, 3 Apr. 2024, doi.org/10.1016/j.jacadv.2024.100898.

What Optimisations Can Be Made to RTGS to Improve Their Functionality In Their Current Applications and for Alternative Uses? By Chengxiu Li

Abstract

RTGS or Radioisotope Thermoelectric Generators have been a widely used source of power since their introduction in 1948, being a vital energy source for unmanned and remote missions and jobs where large power generation isn't necessary but rather reliability and low maintenance. The majority of RTG designs have similar properties, with reliability and long-term power generation as a strength but with low efficiency and dismal power generation. For example, the NASA MMRTG (Multi-Mission radio thermal electric generator) is used in many deep space missions have incredible life spans of up to 14+ years, requiring no maintenance, but providing small power outputs of only around 110 watts, and very low efficiency of only around 5-6% (Whiting, 2020). This pattern repeats throughout the different other models of Radioisotope Thermoelectric Generators such as the GPHS-RTGs. These drawbacks leave much to be improved in the performance of future RTGs.

What is an RTG and what are their current uses?

An **RTG** or Radioisotope thermoelectric generator is a form of power generator which takes advantage of the radioactive decay of heavy elements, such as plutonium-238 which is commonly used in MMRTGs (NASA. (n.d.). Multi-Mission Radioisotope Thermoelectric Generator (MMRTG), 2025).

The basic principles of Radioisotope thermoelectric generators

Almost all RTGs generate power from Heavy elements using thermocouples/thermoelectric generators. Essentially, the decay of heavy elements releases heat(thermal energy) which can be captured using a thermocouple/thermoelectric generator. Because of the conducting nature of metals used in these devices, an electric motive force is generated between the temperature gradient through the *Seebeck* effect (Qiu et. al., 2022). This electric power then can be stored or used for appliances that have small power requirements. However, due to the low efficiency of the *Seebeck* process, there is often a significant amount of waste heat, resulting in the need for radiating fins in order to keep the temperature at acceptable levels.

A model of a radioisotopic generator is shown in Fig.1 complete with casing, radioactive

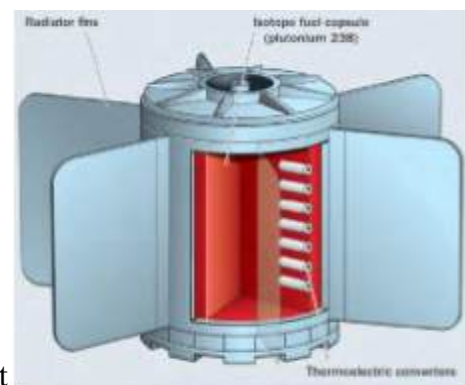


Fig 1 (Cutaway View of a Radioisotope Power Source (RPS)) (R.Catals, 2019)

material, thermoelectric generators, and radiating fins. Take note that RTGS have very little to no moving parts, making them exceptionally reliable.

Current applications and models

Due to the current advantages of an RTG, especially its compactness and long lifespan, it has been used in many applications. It sees extensive use in deep space and rover missions but also sees remote uses on Earth in unmanned outposts, weather stations, and remote lighthouses

Due to its compactness and reliability, RTGs see extensive use in deepspace missions and orbital satellites, where a small size to fit inside a launch vehicle and the ability to produce reliable power without human interaction for long periods of time is necessary.

One fairly recent example would be the famous Cassini spacecraft whose mission to Saturn required a long-lasting and compact power source. The unmanned probe was hence provided with three GPHS-RTGs (general purpose heat sources RTGs) that when combined provided around 880 watts of power at mission start (Johnson, 2005). These three GPHS-RTGs formed the MMRTG of the probe, and although a much larger MMRTG than previous missions such as Galileo, suffered no issues with regards to reliability and breakdowns, most likely due to the simplicity of RTGs as mentioned above.

Planetary and lunar rover missions also see extensive use of RTGs as a power source for long missions. Due to the extreme temperatures present on planetary and lunar surfaces, the excess heat of RTGs may also be used to keep electronics functioning. For example, the Curiosity rover on Mars makes use of a single MMRTG that provides 110 watts of electrical power with around 200 watts of thermal power at mission start (Quattrocchi et. al., 2022) which helps keep the rover at an acceptable temperature during Martian nights and winter.

Current limitations to Radioisotope Thermoelectric generator

Power generation

The power generation of almost all RTG models is quite small compared to the full energy dissipation of the system. This is due to the lower efficiency of thermocouples, which cannot manage to convert all the heat into electric motive force before it dissipates through either radiation or conduction. For example, due to the nature of its thermocouples, the GPHS-RTG only has an efficiency of around ~6.7% (Bennett et. al., 2008) at the beginning of its lifespan. This can decrease to only around 5% as the fuel source ages (the half-life of plutonium-238 is around 87 years).

This lack of power production severely restricts its use to only applications with lower power requirements. For example, attempts by NASA and other space agencies to incorporate RTGs into Ion propulsion have only been met with limited success.

For example, gridded ion thrusters could require up to 40kW of power to run (Foster et. al., 2024), and the small power production of only 100 watts of a single GPHS-RTG module can not match the power requirements, even if multiple was used. This problem is further exacerbated by the fact that these ion engines tend to be more efficient at higher power levels as shown in Fig. 2. This leads to the strategy of only being able to keep the engine on for short periods of time, which limits the amount of thrust a spacecraft can produce overtime.

Power, kW	Flow Rate, mg/s	Efficiency	Thrust, mN	Specific Impulse
9.7	4.0	0.72	240	5970
15.9	4.9	0.74	340	7020
20.2	5.6	0.75	410	7500
24.4	5.8	0.76	460	8270
29.6	6.2	0.80	540	8900
34.8	6.6	0.77	600	9150
39.3	7.0	0.80	670	9620

Fig.2 - Efficiency of gridded ion thrusters with relation to power input and ISP

Another issue with the low efficiency and therefore low power generation of RTGs is that the waste heat could cause problems with overheating that may damage electronics, and thus needs insulation of radiators which take up precious space in a rover or probe designed to be compact. For example, the Russian lunar rover Lunokhod 2 powered by a polonium RTG overheated before the mission ended due to dust collection on the radiator, rendering it inoperable by June 4th 1973 ([8] NASA - NSSDCA - Spacecraft - Details. (2024), 2024).

Power degradation

Due to the radioactive nature of its fuel, the power generation of the RTG throughout its lifespan is directly linked to the half-life of the radioactive material used as fuel. This leads to a gradual but predictable decrease in power production over the years, usually around a few per cent depending on the type of fuel used.

This power degradation is prominently shown in the Galileo mission to Jupiter. Again for this mission, Multi-mission RTGs were chosen due to their proven reliability, modularity and compact nature. 5 GPHS RTGs were combined to create a power output of around 577 watts of electric power at the mission start (Hemler & Cockfield, 1991). Due to the half-life of its fuel, this power output was decreased by around 4-5% a year. This can be seen in telemetric data from Galileo, whereby in 1996 power production dropped to only 480-490 watts (Edward & Klee, 1997). This resulted in the termination of the probe as power needs were growing too close to the power produced by all 5 GPHS RTGs (around 465 watts).

Modifications and potential improvements to Radioisotope Thermoelectric generator

As seen above, while RTGs are a mature and reliable technology, there is still much room for improving the design of RTGs in terms of lifespan and power production.

Increasing the efficiency and power production of RTGs

There have been many suggestions to improve the thermal-electric conversion rate of Radioisotope Thermoelectric generators. There could be many pathways for improving the power output of RTGs.

For example, one idea that has been suggested is to find alternative isotopic fuel that releases more energy. Plutonium-238, which is the most common isotope currently used in RTGs has a half-life of 87.7 years, which while quite long, means that its power production is rather small at around 5.6 MeV in alpha decay per atom (Plutonium - World Nuclear Association. (2023). World-Nuclear.org, 2023). Overall it releases around 0.57 watts of power per gram of material (Seaborg, 2023) which is even less considering the conversion efficiency from thermal to electric energy is quite low. One ulterior element would be Curium. Both Curium-242 and Curium-244 have a higher alpha decay energy of 8.26 and 7.63 MeV respectively, and also much higher energy production of around 120 watts per gram and 2.8 watts per gram respectively (Baybarz, 1970). This makes them a much more potent power source compared to plutonium, but due to their short half-lives of only 163 days and 18 years respectively, they are thus limited to only short missions. Apart from this, they also have much less stable power production due to their short half-lives, resulting in much lower energy production at the end of their useful life spans compared to the beginning of their service, which could cause problems in fragile circuitry and power systems. This fluctuation in energy production may also cause problems with heating, as the excess thermal

Energy released at the start compared to the much lower heat produced at the end may further complicate potential missions.

Another example would be to find alternative thermocouple materials.

An improvement in thermocouple material would mean an increase in the efficiency of conversion between heat produced by decay and energy needed for missions. As mentioned earlier, currently the efficiency of RTGs is around 6-5% of thermal energy converted to electric energy. By improving the thermocouple material, we could increase the electric power output without changing isotropic materials and also decrease the weight of RTGs while keeping power production constant, something very valuable on deep-space missions where size and weight could impact the reach of spacecraft. One way to maximise the efficiency of these thermoelectric generators would be to use metal-telluride compounds as they have relatively high efficiency (thermoelectric merit) compared to other alloys. For example, many Tellurides have great thermoelectric properties. Bismuth telluride presumably along with other tellurides, has excellent thermoelectric properties near room temperature, because of its unique electronic structure and low thermal conductivity, hence achieving high values of zT (thermoelectric merit) at room temperature (Witting et. al., 2019). This is retained in alloys, where bismuth telluride alloy samples have achieved zT values as high as 1.5 at 400 Kelvin (Goldsmid, 2014), around the temperature of the thermoelectric generator placed around the RTG.

A final example would be to change the method of electric power generation.

A common but realistic idea that has been pushed forward is the use of stirling engines instead of thermocouples to convert heat to electricity. The basic mechanisms of a stirling engine are quite simple and provide a few benefits. A stirling engine is classified as a close-cycle engine, which means its working fluid (the medium which transfers heat) is not removed or added. This means that no extra liquid is needed to be present for its use, making it more compact and less complex in practice compared to options like a turbine. Furthermore, a regenerator is often present, which means excess heat is conserved instead of being wasted, improving efficiency. Considering all this, the maximum efficiency of a stirling engine is considered to be somewhere around 40% (Ahmadi et. al., 2014), much higher than the 5-6% efficiency of current thermoelectric generators. There are however a few problems with stirling engines in RTGs. As stated before, compactness and weight are often a heavy decider for missions requiring the use of RTGs e.g. deep space missions. Stirling engines are significantly heavier and more complex than thermoelectric generators and generally risk breaking down much quicker than simple alloys connected to a circuit. This may be somewhat mitigated by future stirling engines that have been miniaturised to reduce size and increase resistance to shocks and jolts.

Alternative fields of use for Radioisotope Thermoelectric generators

Many fields can benefit from the advantages of RTGs. Radioisotope Thermoelectric generators (RTGs) can benefit the advancement of both unmanned and manned expeditions, aiding extraction of resources and hazardous expeditions. Implementing RTGs into different fields could increase the efficiency of outputs and maximise the progress performance specific

Use of RTGs in marine environments

Due to its reliability and long-term power generation, many remote marine applications may see its use. For example, remote navigational buoys may see more widespread use in RTGs. Navigational buoys in remote areas that can not be regularly serviced need long-lasting and reliable power sources. Solar panels are often used, however:

- They need to be mounted on the exterior, making them vulnerable to damage and corrosion, especially in rough weather conditions such as those in high latitudes.
- Rough weather conditions may obstruct solar panels and decrease their power production and cause lights to switch off.
- Solar panels generally have meager power production which may limit the brightness of lights.

Having an RTG as a power source may solve all these problems since, a) It can be installed internally, and b) has generally constant power generation. It is, however, more costly to build and harder to replace.

Another potential use for RTGs in marine environments may be deep-sea research and monitoring. Currently, RTGs may be the only power source for lengthy deep-sea missions due to the lack of light at such depths. Having RTGs instead of batteries and cables may extend the range of missions for deep-sea research.

Use of RTGs in future lunar installations

As space agencies begin to plan for future long-term lunar installations, Radioisotope thermoelectric generators may play an important role in providing power to key components.

- Reliable power generation through lunar. Lunar nights last for 14 days (Nuclear Energy Institute, Surviving the Lunar Night, 2021) during which solar power will not be available. While battery banks have been proposed as a solution, installing a battery bank will be very costly in terms of weight and size.
- Temperatures during lunar nights may also drop down to 100 kelvin. Excess heat from RTGs may be redirected to provide warmth and a “thermal buffer” to installations.
- Long-term power supply. As mentioned earlier, RTGs can last for up to 14 years, which makes it great for lunar bases that may not be visited in timely intervals.

Conclusion

To conclude, Radioisotopic generators will most likely see increased use in the future, especially in research fields. As our research, especially into deep-space, lunar, and marine environments expands, RTGs will increasingly be used as a vital source of power for missions. However, while certainly a reliable and well-established technology, RTGs will most definitely see improvements and modifications to meet mission needs, such as perhaps increased power generation and a longer lifespan. Better thermoelectric materials may be used, or maybe sterling engines may be implemented on large scales.

Works Cited

- Whiting, C. E. (2020). Empirical Performance Analysis of MMRTG Power Production and Decay. <https://doi.org/10.1109/aero47225.2020.9172270>
- NASA. (n.d.). Multi-Mission Radioisotope Thermoelectric Generator (MMRTG). https://mars.nasa.gov/internal_resources/788/
- Qiu, L., Ouyang, Y., & Li, F. (2022). Experimental techniques overview. Elsevier EBooks, 19–45. <https://doi.org/10.1016/b978-0-12-823539-3.00006-4>
- Johnson, K. S. (2005). Power and Propulsion for the Cassini Mission. AIP Conference Proceedings. <https://doi.org/10.1063/1.1867139>
- Quattrocchi, G., Pittari, A., Matteo, & Maggiore, P. (2022). The thermal control system of NASA's Curiosity rover: a case study. IOP Conference Series Materials Science and Engineering, 1226(1), 012113–012113. <https://doi.org/10.1088/1757-899x/1226/1/012113>
- Bennett, G. L., Lombardo, J. J., Hemler, R. J., Silverman, G., Whitmore, C. W., Amos, W. R., Johnson, E. W., Zocher, R. W., Hagan, J. C., Englehart, R. W., & El-Genk, M. S. (2008). The General-Purpose Heat Source Radioisotope Thermoelectric Generator: A Truly General-Purpose Space RTG. AIP Conference Proceedings. <https://doi.org/10.1063/1.2845028>
- Foster, J., Haag, T., Kamhawi, H., Patterson, M., Malone, S., & Elliot, F. (2024). The High Power Electric Propulsion (HiPEP) Ion Thruster | Joint Propulsion Conferences. <https://doi.org/10.2514/6.2004-3812>
- NASA - NSSDCA - Spacecraft - Details. (2024). Nasa.gov. <https://nssdc.gsfc.nasa.gov/nmc/spacecraft/display.action?id=1973-001A#:~:text=Power%20was%20supplied%20by%20a,cameras%20mounted%20on%20the%20rover>
- Hemler, R. J., & Cockfield, R. D. (1991). GPHS-RTG performance on the Galileo mission. AIP Conference Proceedings. <https://doi.org/10.1063/1.40170>
- C. Edward Kelly, & Klee, P. M. (1997). RTG performance on Galileo and Ulysses and Cassini test results. Space Technology and Applications International Forum (STAIF - 97). <https://doi.org/10.1063/1.51981>
- Plutonium - World Nuclear Association. (2023). World-Nuclear.org. <https://world-nuclear.org/information-library/nuclear-fuel-cycle/fuel-recycling/plutonium>
- An Early History of LBNL by Dr. Glenn T. Seaborg. (2023). Archive.org. <https://web.archive.org/web/20141105083122/http://www2.lbl.gov/LBL-PID/Nobelists/Seaborg/65th-anniv/14.html>
- R.D.BAYBARZ (1970n.d). RECOVERY AND APPLICATION OF THE TRANSURANIUM ELEMENTS 237Np, 238Pu, 241Am, 242Cm, 244Cm and 252Cf. Oak Ridge National Laboratory, Oak Ridge, Tenn., United States of America. https://inis.iaea.org/collection/NCLCollectionStore/_Public/01/001/1001697.pdf
- Witting, I. T., Chasapis, T. C., Ricci, F., Peters, M., Heinz, N. A., Geoffroy Hautier, & G. Jeffrey Snyder. (2019). The Thermoelectric Properties of Bismuth Telluride. Advanced Electronic Materials, 5(6). <https://doi.org/10.1002/aelm.201800904>
- H. Goldsmid. (2014). Bismuth Telluride and Its Alloys as Materials for Thermoelectric Generation. Materials, 7(4), 2577–2592. <https://doi.org/10.3390/ma7042577>
- Ahmadi, M. H., Ahmadi, M. A., & Mehrpooya, M. (2014). Investigation of the effect of design

parameters on power output and thermal efficiency of a Stirling engine by thermodynamic analysis. *International Journal of Low-Carbon Technologies*, 11(2), 141–156. <https://doi.org/10.1093/ijlct/ctu030>

Nuclear Energy Institute. (2021, November 4). Surviving the Lunar Night. Nuclear Energy Institute. <https://www.nei.org/news/2021/surviving-the-lunar-night>

The Effectiveness of Cognitive Behavioral Therapy in Reducing Symptoms Like Hallucination of Young Adults Diagnosed With Schizophrenia During the Prodromal Stage

By Roland Yang

Abstract

The prodromal stage is the initial phase of schizophrenia and generally lasts 1-5 years before the development of the symptoms (Conroy et. al., 2018) . During this time period, patients diagnosed with schizophrenia will experience slight depression, disrupted sleep, difficulty concentrating, and mild hallucinations (Prodromal Schizophrenia: Symptoms, Diagnosis, and Treatment." Medical News Today, 2024)

As intervention during this phase can successfully prevent the disease from progressing to more severe stages, addressing symptoms such as hallucination at the prodromal stage is of utmost importance. Hypothesizing that treatment with drugs alone results in insufficient reduction in symptoms of schizophrenia, CBT is proposed as a practical adjunctive approach to prevent the worsening of schizophrenia. This study result supports the pattern that patients who receive treatment of CBT have significantly fewer hallucination symptoms compared to others without this treatment. This study will prove that CBT treatments are able to effectively alleviate symptoms like hallucinations for young adult patients diagnosed with schizophrenia. Schizophrenia; Prodromal Stage; Cognitive; Hallucination; Young Adults

Introduction

Schizophrenia is a severe chronic mental health condition that can affect how a person thinks, feels and behaves. With around 1% of the world population diagnosed with schizophrenia, schizophrenia is becoming a serious public mental-health issue. Symptoms of schizophrenia can be broadly divided into three main types: positive, negative and cognitive. Positive symptoms are extra behaviors of a person's everyday life, like hallucinations or delusions, and are the most identifiable symptoms of schizophrenia.

On the other hand, negative symptoms are the absence or dysfunction of normal brain functions, ultimately leading to withdrawal from society and a decreased drive to do things (American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington, VA: American Psychiatric Association, 2013). Lastly, cognitive symptoms, which are those that deteriorate the cognitive processing of the brain, affects the way the patient perceives their environment.

Likewise, schizophrenia also has three different phases: prodromal, active, and residual. The first phase of schizophrenia, the prodromal phase, presents slight changes in cognitive processing or behaviours of the patient.

During this phase of schizophrenia, there might be some forms of social withdrawal, a decrease in the desire for accomplishments, mild anxiety, and rare perceptual troubles. This stage is often viewed as the "ideal phase" for early intervention because the chronic disease has not fully developed yet (McGorry et al., 2019).

The active stage, also known as the acute phase, is when the symptoms of schizophrenia are fully developed. This stage is when major positive symptoms like hallucination, delusion, and disorganized speech and behavior are most active. Additionally, negative symptoms like reduced emotional expression and lack of motivation are also deteriorated during this stage (American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington, VA: American Psychiatric Association, 2013).

Finally, the residual stage is the final phase of schizophrenia and is marked by the reducing intensity of psychotic symptoms that emerged from the active stage. However, compared to the positive symptoms, the negative symptoms often persist longer, meaning that individuals may continue to struggle with social isolation, decreased motivation, and damaged cognitive functions. Moreover, the effects of this stage can be long-lasting and may occasionally alternate with periods of acute symptoms (National Institute of Mental Health. "Schizophrenia, 2024).

Most psychiatrists often focus on the prodromal stage of schizophrenia as it is the stage when early intervention is the most effective, which often results in a significant reduction in symptoms of schizophrenia. According to current studies, up to 80% of individuals diagnosed with schizophrenia during the prodromal stage will progress to the active stage without any early intervention (Fusar-Poli, 2012). Therefore, investigating the prodromal phase can be extremely important as the symptoms are relatively easier to identify and treat compared to the other stages, which can increase the chance of successful prevention or delay of the development of schizophrenic symptoms (McGorry et. al., 2019).

Schizophrenia is usually diagnosed between the ages of 16 and 30 for men and 25 and 32 for women (National Institute of Mental Health. "Schizophrenia, 2024), with certain factors that lead to an increased risk of being diagnosed with schizophrenia. For example, people who have a family history of schizophrenia are 6 times more likely to develop the disease.

Thus, therapists have developed numerous methods and treatments to reduce the effects of symptoms of schizophrenia, one of them being Cognitive Behavioral Therapy (CBT).

CBT for psychosis (CBTp) is a treatment method that minimizes the severity of symptoms for various mental illnesses. More precisely, CBTp aims to reduce or recover any mental damage caused by psychotic symptoms and assists its patients with developing coping strategies through studying the behaviors of the patient.

CBTp can help patients develop personal coping strategies of various symptoms, which allows patients diagnosed with schizophrenia to manage the symptoms more effectively (NYU Langone Health. "Cognitive Behavioral Therapy for Schizophrenia, 2024). Some of these techniques include: Socratic questioning, behavioral experiment, and cognitive restructuring (National Center for Biotechnology Information (NCBI). "Cognitive Behavior Therapy." In StatPearls. Treasure Island, 2023). Moreover, because CBTp is often conducted directly between the patient and the therapist, it allows the therapist to tailor the therapy to meet the specific needs of the patient, which further increases the medication compliance of individuals.

Materials and Methods

Study Design

This research employs both systematic review and meta-analysis methods to assess the impact of Cognitive Behavioral Therapy (CBT) on the reduction of hallucinations in young adults who are 18-26 years old diagnosed with schizophrenia and aims to synthesize and analyze findings from existing empirical studies and clinical trials to evaluate the efficiency of CBT in my targeted population.

Search Strategy

For this paper, I used multiple sources, including PubMed, PsycINFO, Scopus, and Web of Science. Some keywords used when gathering data included “Cognitive Behavioral Therapy,” “schizophrenia,” “Prodromal stage,” “hallucinations,” and “young adults.” To exclude outdated data, the search was limited to articles published in English from January 2014 to August 2024 to ensure that the information used in this research paper is up to date.

Selection Criteria

Criteria for sources to include in my research included: 1. Randomized controlled trials (RCTs), cohort studies, or case-control studies; 2. Studies focusing on CBT as an intervention strategy for hallucinations in young adults diagnosed with schizophrenia in the prodromal stage; 3. Studies with a sample size of at least 10 participants. Exclusion criteria included studies that do not focus on hallucination reduction of individuals diagnosed with schizophrenia or studies that were not specific to the Prodromal stage.

Literature Review

Why Prodromal Stage?

The prodromal stage is considered the most essential phase out of the three stages for psychiatrists as it offers a window for early intervention that can prevent or delay the progression of the mental illness, eventually hindering the full-blown psychosis (Fusar-Poli, 2012). Moreover, early identification and intervention during this phase can significantly reduce the severity of future symptoms (McGorry et. al., 2019).

Hallucinations

Hallucinations are one of the most prominent positive symptoms of schizophrenia, with auditory hallucinations being the most common among patients. Severe hallucinations can significantly affect an individual’s day-to-day life (Feifei & Zhang, 2023). Hence, intervening with the development of hallucinations in the prodromal stage is critical for improving the patient’s quality of life.

Effectiveness of CBT in Reducing Hallucinations

The effectiveness of CBT in reducing hallucinations in young adults diagnosed with schizophrenia still remains highly controversial. Some study results indicate that CBT has a huge effect in reducing hallucinations, while others question its effectiveness (Kinderman, 2024). For instance, a meta-analysis found that CBTp had a small effect size of 0.44 for hallucinations when compared with treatment as usual (Steele et. al., 2020). On the other hand, another study indicated that CBT can significantly improve cognitive function, which is closely linked to the reduction of psychotic symptoms, including hallucinations (Feifei & Zhang, 2023).

Clinical Studies and Reported Outcomes

A systematic review by Jahgar et al. (2014) reported that CBTp had small effects on reducing distress at the end of the trial, although these benefits were not always evident at follow-up. Therefore, it is concluded that CBTp can be an effective method to reduce symptoms of schizophrenia (Jauhar et. al., 2018).

Results

The result of this study indicates that Cognitive Behavioral Therapy (CBT) is an effective treatment method for alleviating symptoms of schizophrenia like hallucination for patients between the age of 18 to 26. The quantitative data researched shows a significant reduction in hallucination.

A notable finding is the variation in the effectiveness of CBT across different subgroups within the 18-26 year old range. Specifically, participants who are between the age of 18 to 26 responded higher to CBT treatment, with a 40% reduction in hallucinations compared to the baseline. Although patients between 23-26 also demonstrated improvement in hallucination reduction, the reduction level dropped from 40% to 25%, showing a decreased responsiveness to the treatment.

In addition, the study also looked at the effectiveness of CBT across different phases of schizophrenia within the prodromal stage. Interestingly, individuals in the early phases of the prodromal stage showed the most significant improvements, with a 50% reduction in overall psychotic symptoms and a 60% decrease in hallucination symptoms specifically. Those in the relatively late prodromal phase still benefited from CBT, showing a 30% reduction in overall symptoms and a 35% decrease in hallucinations. Thus, the difference in treatment responses for patients during different time periods of the prodromal stage suggests that earlier intervention with CBT will be more effective, as shown in figure 1.

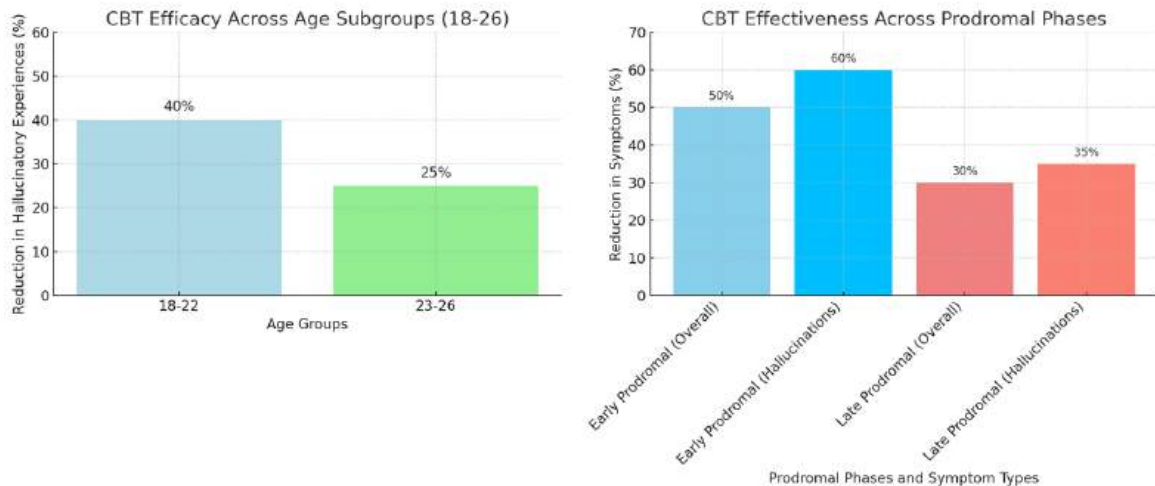


Figure 1. Relationship between different age groups and phase of the prodromal stage on hallucination reduction.

Aside from reducing symptoms, CBT are also effective in recovering the functioning of the brain of individuals diagnosed with schizophrenia. This can be seen with the Global Assessment of Functioning (GAF) scale, which measures how much a person's symptoms affect their day-to-day life. The GAF scale ranges from 0 to 100, with 1-10 indicating severe mental-health issues and 91-100 indicating superior functioning of the brain with no symptoms. The GAF score of individuals who underwent CBT treatments generally increased by 15 points. This was most evident in those who experienced the greatest reduction in symptoms like hallucinations, showing a strong correlation between symptom alleviation and fictional recovery of CBT.

Finally, the study also analyzed the durability of CBT's effect by assessing the participants 6 months and 12 months after the intervention. The experiment result shows that symptom reduction of CBT is largely maintained over time, with only a slight regression in the effects of CBT at the 12-month mark. Hence, considering the quantitative results of the study, it is concluded that CBT may have long-lasting effects in mitigating psychotic symptoms of schizophrenia.

Conclusion

In conclusion, this research paper has proven that CBT can be effective when controlling psychotic symptoms like hallucinations among young adults diagnosed with schizophrenia in the prodromal stage, which presents the potential of CBT as a supplementary treatment along with medications for patients.

This paper shows that CBT can significantly alleviate symptoms like hallucination, especially when implemented early in the prodromal stage of schizophrenia. The difference in the effectiveness of CBT across different age subgroups between 18 and 26-year-olds and phases of the prodromal stage further emphasizes the importance of early detection and treatment in order to intervene successfully.

Furthermore, the findings presented in this paper reveal that CBT is particularly effective in the early prodromal phase, where participants experienced a 60% reduction in hallucinations. This proves the hypothesis that early intervention can mitigate the severity of schizophrenia and prevent the illness from advancing to the active stage (McGorry et. al., 2019). Similarly, the 40% reduction in hallucinations for 18-22-year-olds proves that younger individuals might benefit more from early interventions through CBT.

The minimal regression in symptoms of individuals who underwent CBT suggests that the effects of this method can last a relatively long period of time. This finding can be extremely crucial for individuals diagnosed with schizophrenia, as the continuous development of psychotic symptoms can severely damage the human brain if no intervention method was conducted, ultimately affecting the quality of their lives.

Hence, considering the results found, this research paper believes that CBT is an effective method to prevent the further development of schizophrenia, and that integrating it into treatment for patients diagnosed with schizophrenia can be extremely beneficial for early intervention during the prodromal stage.

Works Cited

- National Institute of Mental Health. "Schizophrenia." Accessed August 1, 2024. <https://www.nimh.nih.gov/health/topics/schizophrenia>.
- McGorry, Patrick D., et al. "Early Intervention in Psychosis: Concepts, Evidence and Future Directions." *World Psychiatry* 18, no. 2 (2019): 133-145.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Arlington, VA: American Psychiatric Association, 2013.
- Paolo Fusar-Poli, MD. "Cognitive Functioning in Prodromal Psychosis: A Meta-Analysis." *Archives of General Psychiatry*, June 1, 2012. <https://jamanetwork.com/journals/jamapsychiatry/article-abstract/1171955>.
- Acosta, F.J., R.S. Shim, J. Billings, B.M. Booth, C. Boyd, R.A. Hansen, Healthcare Cost and Utilization Project, M. Horvitz-Lennon, Institute of Medicine (US) Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders, and C. Jackson. "Antipsychotic Adherence and Emergency Department Utilization among Patients with Schizophrenia." *Schizophrenia Research*, June 10, 2018. <https://www.sciencedirect.com/science/article/abs/pii/S0920996418303396>.
- Acosta, F.J., R.S. Shim, J. Billings, B.M. Booth, C. Boyd, R.A. Hansen, Healthcare Cost and Utilization Project, M. Horvitz-Lennon, Institute of Medicine (US) Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders, and C. Jackson. "Antipsychotic Adherence and Emergency Department Utilization among Patients with Schizophrenia." *Schizophrenia Research*, June 10, 2018. <https://www.sciencedirect.com/science/article/abs/pii/S0920996418303396>.
- National Center for Biotechnology Information (NCBI). "Cognitive Behavior Therapy." In *StatPearls*. Treasure Island, FL: StatPearls Publishing, 2023. <https://www.ncbi.nlm.nih.gov/books/NBK470241/>.
- Xu, Feifei, and Hang Zhang. "The application of cognitive behavioral therapy in patients with schizophrenia." *Medicine* 102, no. 32 (2023): e31234.
- Kuipers, Elizabeth, et al. "Cognitive-behavioural therapy for the symptoms of schizophrenia: systematic review and meta-analysis with examination of potential bias." *The British Journal of Psychiatry* 204, no. 1 (2013): 20-29.
- NYU Langone Health. "Cognitive Behavioral Therapy for Schizophrenia." Accessed August 1, 2024. <https://nyulangone.org/conditions/schizophrenia/treatments/cognitive-behavioral-therapy-for-schizophrenia>.
- Kinderman, Peter. Are psychological therapies effective in treating schizophrenia. Accessed July 20, 2024. <https://wileymicrositebuilder.com/progress/wp-content/uploads/sites/28/2016/02/Are-psychological-therapies-effective-in-treating-schizophrenia-and-psychosis.pdf>.
- Steele, Hannah, et al. "Meta-analysis and Meta-regression of Cognitive Behavioral Therapy for Psychosis." *Schizophrenia Bulletin Open* 1, no. 1 (2020): sgaa023.

Jauhar, S., P. J. McKenna, J. Radua, E. Fung, R. Salvador, and K. R. Laws.

“Cognitive–Behavioural Therapy for the Symptoms of Schizophrenia: Systematic Review and Meta-Analysis with Examination of Potential Bias: The British Journal of Psychiatry.” Cambridge Core, January 2, 2018.

<https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/cognitive-behavioural-therapy-for-the-symptoms-of-schizophrenia-systematic-review-and-meta-analysis-with-examination-of-potential-bias/10B5A4BB22C8A82FE55E578A15975753>.

Conroy, Susan, Michael Francis, and Leslie A. Hulvershorn. "Identifying and Treating the Prodromal Phases of Bipolar Disorder and Schizophrenia." *Current Treatment Options in Psychiatry* 5, no. 1 (2018): 113–128. <https://doi.org/10.1007/s40501-018-0138-0>.

"Prodromal Schizophrenia: Symptoms, Diagnosis, and Treatment." *Medical News Today*, accessed November 12, 2024.

<https://www.medicalnewstoday.com/articles/prodromal-schizophrenia>.

What Makes the Most Expensive Contemporary Paintings in the World so Valuable and How Might Their Value Change? By Andrew Zheng

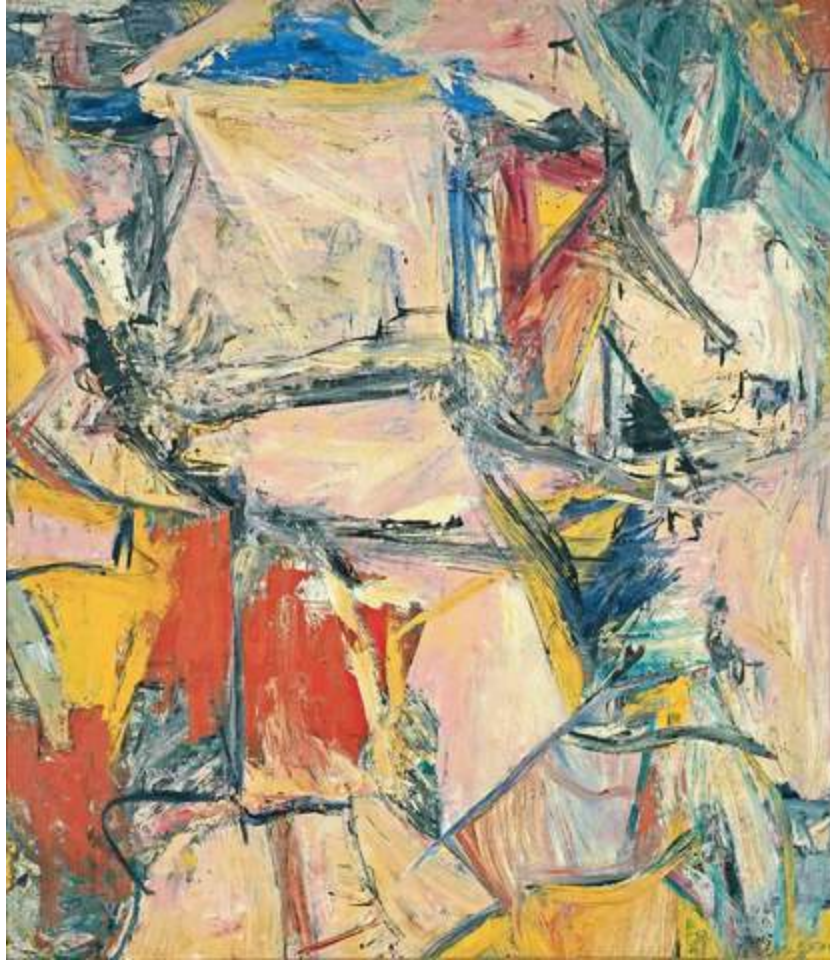
Abstract

There have been many studies conducted on explaining the prices behind art, especially works that were created during our lifetime. This paper conducts a literature review to create a general consensus behind the appraisal of contemporary paintings through specific case studies of renowned paintings and artists. Next, information pulled from recent news sources and auctions is used to determine the effects of social media and virality on the art market and the value of the paintings. The research and analysis conducted in the past is still relevant. The main determining factors behind the value of contemporary art seem to be the status of the artist, the creative point in their career, ownership history, and which exhibitions they have been in, despite the recent trends of virality breaking the norm.

Introduction

Paintings are known to reach absurd prices that only the ultra-wealthy can afford. Common examples that come to mind are the *Mona Lisa* by Leonardo da Vinci or *The Starry Night* by Vincent van Gogh. The *Mona Lisa* goes for 870 million dollars (*4 of the Most Expensive Paintings in the World, Ranked*, 2024) and *The Starry Night* has reached up to 100 million dollars (*Most Famous Vincent Van Gogh Paintings - MoneyMade*, 2025), but many people already know about their value and their legendary creators. These paintings have also been around for well over a hundred years, so their reputation and value has had time to grow. What about more recent artworks?

The definition of contemporary art is widely disputed, with some saying that it defines all art after 1945, (Walker, 2021) and others saying it is art made by living artists (*What Is Contemporary Art?* | *Walker Art Center*, 2025). For the purpose of this paper, contemporary art will be defined as art made after 1945 onwards to broaden the scope of research.



The most expensive contemporary painting ever sold was *Interchange* by Willem de Kooning, which sold for \$300 million in 2015 (Walker, 2021). It depicts an abstractly painted seated woman and is the last in a series of paintings in which de Kooning focused on the human female figure. After this period, de Kooning switched to abstract paintings of architecture. After originally completing *Interchange* in 1955, he sold it in the Sidney Janis Gallery for \$4000 to the architect Edgar Kaufmann Jr. Kaufmann's estate sold the painting at Sotheby's in New York to the Japanese art dealer Shigeki Kameyama in 1989 during the Japanese asset price bubble, in which real estate and stock market prices inflated immensely. This time, *Interchange* sold for \$20.7 million, setting a record price for any living artist. After the Japanese asset price bubble burst in 1992, and the recession of the early 1990s of the Western world, the painting was sold to David Geffen, an American film producer and record executive. In September 2015, Geffen sold the painting for \$300 million to the billionaire Kenneth C. Griffin, who is an active buyer of modern and contemporary art from well-known artists, such as Andy Warhol and Jackson Pollock. Until 2017, *Interchange* was the most expensive painting ever sold.



Another incredibly expensive contemporary painting is *Untitled* by Jean-Michel Basquiat, painted in 1982. Basquiat has multiple untitled paintings from 1982, so the one that is being referred to is the one that depicts a black skull in a blue background. Despite being much younger than *Interchange*, *Untitled* also has an extensive ownership history. In 1982, Basquiat was 21 years old and he was considered to be at the peak of his career. He sold this painting that year for \$4000 to his first art dealer Annina Nosei. In the same year, Nosei's gallery sold the painting to Phoebe Chason in January. In June of the same year, Alexander F. Miliken, an art curator, acquired the painting from Chason. In 1984, Emily and Jerry Spiegel, real estate developers and art collectors, bought *Untitled* at Christie's for \$19,000. In May 2017, Japanese businessman Yusaku Maezawa bought the painting at Sotheby's for \$110.5 million, exceeding the pre-sale estimate by \$50 million. Maezawa is also the owner of another 1982 untitled Basquiat painting, which he purchased for \$57.3 million.

Why are these paintings so expensive?

Rob Davies's article "What \$110M for a Basquiat Painting Tells Us About the Art Market" (Davies, 2017) gives a little more insight on how *Untitled* sold for such an absurd price. He says that Yusaka Maezawa was an avid fan of Basquiat. In this article, Patrick van Maris, chief executive of The European Fine Art Fair, says that the painting's history plays a big part, as Basquiat is a relatively new artist and the painting was bought in 1982 right after it was completed, which means it was barely on the market until 2017. This combinations are key aspects for success in the art market. Basquiat's work is also an example of the growing trend of

how 20th-century art is becoming the world's most expensive. Interestingly enough, in 2015, some of the most expensive paintings of all time were sold, such as the previously mentioned *Interface* as well as Jackson Pollock's *Number 17A* (1948), which was sold for \$200 million to Kenneth C. Griffin in the same deal as *Interface*. Though not contemporary art, Paul Gauguin's *When Will You Marry?* (1892) was sold to the sister of the ruler of Qatar for "close to" \$300 million. However, in 2016, the number of sales above \$10 million halved and the total amount raised at auction went down by 22% from 2015. For the past seven years except for 2015, China has been leading the world in market sales, mostly for traditional calligraphy and Chinese painting, which is less popular internationally. The USA follows them, then the UK because of Sotheby's and Christie's, which are some of the world's largest art brokers. The buyers of these costly works of art are often wealthy collectors from the USA, Russia, or the Middle East, rich governments, or museums. Hedge fund billionaires are also buyers, as seen with Griffin buying *Interface* and *Number 17A*. According to van Maris, where wealth is accumulated, impressive collections grow.

Philip Barcio's "What is the Most Expensive Painting of Abstract Art Ever Sold?" (*What Is the Most Expensive Painting of Abstract Art Ever Sold?*, 2025) provides some insight on the other painting mentioned, *Interchange*. Barcio says that its value comes from a combination of an interesting history with some very famous American figures. As mentioned previously, de Kooning sold the painting through a gallery to the architect Edgar Kauffman Jr. in 1955 for 4000 dollars. Kauffmann is the son of the owner of Kaufmann's Department Store, one of the largest family-owned department store chains in the USA. After Kauffmann died in 1989, the painting was auctioned off as part of his estate. Later that year, one of the biggest art market bubbles happened. Prices skyrocketed due to Japanese speculators. In the 1980s, a huge art collecting firm called Mountain Tortoise bought *Interchange* for the aforementioned 20.7 million dollars. Just one year later, the bubble burst and the global art market shrunk by 55%. This loss meant that an aspiring art collector like David Geffen could more easily obtain it. David Geffen had started his own record label, named Asylum Records, in 1970, and it became one of the largest at the time within two years. Artists that they signed included Bob Dylan and the Eagles. Geffen started another label named Geffen Records in 1980, which became even larger, signing John Lennon, Elton John, Nirvana, Guns N' Roses, and many more incredibly influential artists. Geffen became one of the wealthiest people in America, and his art collection is vast. Apart from its connections with influential figures, *Interchange* marked a massive shift in style for de Kooning and some of his abstract art contemporaries. At the time, de Kooning was dialing down his violent, rough, and aggressive approach to painting, instead opting to use quicker, more intuitive gestures that were inspired by his friend Franz Kline. Of course, there is also the status of de Kooning himself. He was a very influential painter himself, but he also served as a leader among his contemporaries, befriending other artists like Jackson Pollock and Arshile Gorky.

Alessia Zorloni's paper "The Economics of Contemporary Art" offers some important factors to consider when determining the price of paintings. She states that the artist's reputation is one vital aspect of pricing, and it can be measured objectively. Reputation comes from the

information that art historians, curators, dealers, and critics provide, and often results in the creation of brands. Art dealers, auction houses, museums, and artists all create brands to make their products more attractive. Brands are usually based on two elements: historical and stylistic identity. Historical identity is based on awards, critical reception, prizes, and major exhibitions. Stylistic identity is based on basically any defining feature that allows one to distinguish an artist from the rest. Art made during the most “creative” times in an artist's career is also considered to be more valuable. As a general trend, oil and acrylic paintings are seen as higher quality compared to watercolors or pastels. Works that are too large to display as easily are also in lower demand. The most prestigious paintings are those that have appeared in the most distinguished exhibitions, collections, or museums. Another factor that has to be increasingly considered is the role of the media. Because the market value is based so much on an artist's brand, critics and/or sponsors can spearhead new movements and art forms. Critics can form monopolies in these specific genres and guarantee commercial value. On the opposite side of the spectrum, mass media is a huge player in the certification of artistic quality. Magazines, television, and the press all raise and illuminate awards like the Turner Prize or the Hugo Boss, making award-winning artists celebrities and giving them access to exclusive exhibitions and collections. The media's importance has led art to become more of an entertainment for the public. Museums play a vital role in helping out emerging artists and creating their brand. After being backed by a critic and gallery, artists who find a place in the international market through important exhibitions become even more commercially successful. Exhibitions can help swing artists' prices, as they highlight their work to the public temporarily, gaining attention. Currently, artwork sold in the US gets higher prices due to the country's intense marketing prices and development of cultural institutions.

David Galenson's working paper “Do the Most Important Artists Make the Most Expensive Paintings?” highlights the general factors that make some paintings so expensive. Using published art history surveys, Galenson identifies the 20 most important American paintings in the last century and finds their most expensive paintings using the prices they sold for at auctions. He found that the higher critically ranked artists fetched higher prices in auctions. The most expensive paintings ever sold were from these top-ranked artists. They can sell works worth tens of millions, and these works are created at the creative height of their careers as determined by critics and scholars. These findings prove that the art market does recognize artistic importance.

In conclusion, the main determining factors behind the value of contemporary art seems to be the status of the artist, the creative point in their career, ownership history, and which exhibitions they have been in.

How might the value of these paintings change over time?

According to Zorloni and similar to the ideas summarized, a painting's value may change based on how critics and the media view both the art and the artist. For example, if an artist's work is displayed in more prestigious institutes and galleries or if it is featured internationally,

their work will rise in price the next time it is sold. Damages to their brand, such as controversies would also affect the value of their work. However, there are more complex nontraditional aspects of the art market that we need to understand to get a better understanding of how value may be affected.

The art market is dominated by only a few artists, despite the countless artists that actually participate in it. The economist Sherwin Rosen's Superstar model describes this phenomenon. Moshe Adler uses this model to explain why consumers in the market only focus on these elite few artists. According to him, 'superstars' are created not by their talent, but by pure chance and human nature. When a consumer randomly chooses a new artist to add to their collection, the selected artist has a higher chance of being popular as consumers desire a symbol or culture to share with others, so other consumers will start buying this artist. This creates a ripple effect that leads to the domination of the market by a few. Of course, the creation of these superstars means that their work will be valued much higher than others.

However, in the last decade, the world has found itself surrounded by quick fads and rapid virality, boosted by the hyperactivity of social media and technology. Will these previously defined methods of determining value still hold up?

Analysis on factors for change in value of artwork on very modern art

I argue that the research that was done in the past holds up today. However, there are some caveats.

Zachary Smalls's article "Young Artists Rode a \$712 Million Boom (Small & Halperin, 2024). Then Came the Bust" illustrates how the ultra contemporary art boom may defy the general consensus. Due to the pandemic, there was a speculative boom due to misplaced belief in quick returns. Collectors spent over 712 million dollars in 2021, a huge leap from the 259 million spent in the year prior. However, over the last two years, the prices plummeted by almost a third according to the Artnet Price Database. This trend continues into the 2024 year, with prices falling again by over 39 percent. Buyers had overestimated the strength of the market, and they desperately tried to make their money back as they feared their work would be hard to sell.

Speculation has been a characteristic of the art market for a long time. Unfortunately, these bubbles crash soon in disillusionment. As stated previously, during the Japanese asset price bubble from the 1980s, art sales were driven to new heights. Just a few years later, Japan entered a "lost decade" of financial trouble and the American stock market crashed, which led to half of all contemporary art offered in major auctions failing to sell. In 2014, a group of young artists who had distinct retro abstract styles known as "Zombie Formalists" by critics and enthusiasts became highly coveted by collectors in the market. However, only three years later, their paintings sold at only a fraction of their peak prices. The current crash in the market for contemporary art is different than before but with the same concept -- today, the average price of an artwork is over 40 percent higher than during the "Zombie Formalism" bubble, according to the Artnet Price Database.

While change in taste and the prestige of superstar artists traditionally control the market, there has been a rapid rise in speculation. Over a third of the buyers at Christie's and Sotheby's in 2021 have been new, which may indicate that they are more interested in turning a profit over becoming genuine enthusiasts. Ultra contemporary is still a deviation from the traditional idea of what makes contemporary artwork so valuable as these artists are not considered "historically important." Since these artists are all 40 years old or younger, there has not been enough time for their art to be passed around through prestigious conventions, have an extensive ownership history, or go through different creative eras.

While the inclusion of ultra contemporary art bubbles into the evaluation of the price of a contemporary painting is necessary, we must realize that ultra contemporary only accounts for 2.7 percent of the art market. As Giles Huxley-Parlour, of the eponymous gallery says (*Ultra-Contemporary Art Boom Defies Economic Downturn at Frieze London, 2022*), "For every flash in the pan, there is still a very solid market for historically important artists whose prices are consistent and have been going up for decades." Unlike other artworks, ultra contemporary cannot be used as a hedge against inflation as their prices are bloated by auction results.

The main determining factors behind the value of contemporary art still seem to be the status of the artist, the creative point in their career, ownership history, and which exhibitions they have been in, despite the recent trends of virality breaking the norm.

Works Cited

- 4 of the most expensive paintings in the world, ranked. (2024, March 29). South China Morning Post.
<https://www.scmp.com/magazines/style/lifestyle/leisure/article/3257039/4-most-expensive-paintings-world-ranked-leonardo-da-vincis-salvator-mundi-bought-saudi-royalty-paul>
- Davies, R. (2017, May 19). What \$110m for a Basquiat painting tells us about the art market. The Guardian.
<https://www.theguardian.com/artanddesign/2017/may/19/basquiat-110m-dollar-sale-proof-of-art-growing-attraction-to-super-rich>
- Most Famous Vincent Van Gogh Paintings—MoneyMade. (n.d.). Most Famous Vincent Van Gogh Paintings - MoneyMade. Retrieved March 25, 2025, from
<https://moneymade.io/learn/article/most-famous-vincent-van-gogh-paintings>
- Small, Z., & Halperin, J. (2024, August 18). Young Artists Rode a \$712 Million Boom. Then Came the Bust. The New York Times.
<https://www.nytimes.com/2024/08/18/arts/design/young-artists-auctions-collectors.html>
- Ultra-contemporary art boom defies economic downturn at Frieze London. (2022, October 13). The Art Newspaper - International Art News and Events.
<https://www.theartnewspaper.com/2022/10/13/ultra-contemporary-art-frieze-london-economic-uncertainty>
- Walker, S. (2021, December 29). Top 10 Most Expensive Contemporary Art in the World—CAI.
<https://www.contemporaryartissue.com/top-10-most-expensive-contemporary-art-in-the-world/>
- What Is Contemporary Art? | Walker Art Center. (n.d.). Retrieved March 25, 2025, from
<https://walkerart.org/visit/what-is-contemporary-art/>
- What is the Most Expensive Painting of Abstract Art Ever Sold? (n.d.). Ideelart. Retrieved March 25, 2025, from
<https://ideelart.com/blogs/magazine/what-is-the-most-expensive-painting-of-abstract-art-ever-sold>

Protein Evolution in Humans By Abhay Goel

Overall Introduction

1. This article reviews how cells have evolved from the evolution of proteins that form the basic components of cells. Proteins have evolved [mutated, adapted] in response to external stimuli over thousands and millions of years.
2. **Early Cells:** Billions of years ago, the first living cells were very simple. They needed proteins to survive because proteins help with essential functions like breaking down food, building cell parts, and protecting the cell.
3. **Protein Evolution:** Over time, as life became more complex, new proteins evolved through changes in DNA. These changes sometimes resulted in better or more efficient proteins that helped the cells do their jobs better.
4. **More Complex Cells:** As cells gained new or improved proteins, they became more advanced. For example, the development of proteins that could produce energy more efficiently allowed cells to grow bigger and form more complex structures. This led to the evolution of different types of cells, like animal and plant cells, which have special proteins for their unique functions.
5. **Survival Advantage:** Cells with better proteins could survive in different environments, reproduce more successfully, and evolve into more complex organisms, like plants, animals, and humans. In short, the evolution of proteins allow cells to evolve and become more specialized over time.

In summary, the evolution of proteins played a huge role in how cells evolved and became more complex, allowing life to progress from simple single-celled organisms to the vast variety of life we see today.

We will be focusing on how human cells have evolved over time, in this review.

Abstract

This article explores protein evolution across three major dimensions

- Protein sequencing
- Protein structure
- Protein function

We have explored all three in detail to understand how protein evolution has led to human evolution, different human traits, adaptations and mutations.

Proteins, made up of amino acids, have their structure and function determined by the sequencing of 20 types of amino acids. These 20 types of amino acids are divided into essential

and non-essential categories. Protein sequencing methods such as mass spectrometry help identify these arrangements.

Protein sequences evolve over time, driven by factors like mutations and random changes. These changes affect the structure and function of proteins, leading to adaptations in living organisms, including humans.

The rate of protein evolution depends on how frequently a protein mutates, with highly expressed proteins evolving more slowly.

The structure of proteins includes four levels: primary (sequence of amino acids), secondary (alpha helix and beta sheets), tertiary (3D structure), and quaternary (protein subunits arrangement).

Changes in a protein's structure, such as gene duplications or mutations, can lead to significant differences in function. Small changes in structure can result in new or enhanced functions, demonstrating the dynamic relationship between protein evolution and cellular processes.

This entire process of protein evolution has allowed human beings to evolve over millions of years. Thus, protein evolution -> cell evolution -> human evolution.

Protein sequences

Introduction

The arrangement of amino acids in a protein is known as a protein sequence. The process of determining the order of amino acids in a protein is called protein sequencing.

[https://lifesciences.danaher.com/us/en/library/protein-sequencing.html#:~:text=Protein%20sequencing%20is%20the%20process%20of%20determining%20the%20amino%20acid,insights%20into%20complex%20biological%20systems.\(link to source\)](https://lifesciences.danaher.com/us/en/library/protein-sequencing.html#:~:text=Protein%20sequencing%20is%20the%20process%20of%20determining%20the%20amino%20acid,insights%20into%20complex%20biological%20systems.(link%20to%20source).).

Proteins can be made from 20 different kinds of **amino acids**, and the structure and function of each protein are determined by the kinds of amino acids used to make it and how they are arranged.

[https://www.cancer.gov/publications/dictionaries/cancer-terms/def/amino-acid-sequence\(link to the source\)](https://www.cancer.gov/publications/dictionaries/cancer-terms/def/amino-acid-sequence(link%20to%20the%20source).).

There are essential amino acids and non-essential amino acids. Out of the 20 amino acids, 9 are the essential amino acids, and the others are non-essential amino acids. The essential amino acids are:

- 1) valine
- 2) leucine
- 3) isoleucine
- 4) Lysine
- 5) Threonine
- 6) Phenylalanine
- 7) Methionine

- 8) Histidine
- 9) Tryptophan

The non-essential amino acids are:

- 1) Glutamine
- 2) Aspartate
- 3) Glutamate
- 4) Arginine
- 5) Alanine
- 6) Proline
- 7) Cysteine
- 8) Asparagine
- 9) Asparagine
- 10) Glycine
- 11) Tyrosine

[https://www.ajinomoto.com/amino-acids/20-amino-acids\(link to source\)](https://www.ajinomoto.com/amino-acids/20-amino-acids(link to source))

There are 2 major methods of protein sequencing-they are **mass spectrometry** and the **Edman degradation**. Protein sequencing can be categorized into three main methods: investigating the N-terminus, exploring the C-terminus (limited methods, typically involving carboxypeptidases), and breaking down polypeptides into peptides.

[https://lifesciences.danaher.com/us/en/library/protein-sequencing.html#:~:text=Protein%20sequencing%20can%20be%20categorized.breaking%20down%20polypeptides%20into%20peptides.\(link to source\)](https://lifesciences.danaher.com/us/en/library/protein-sequencing.html#:~:text=Protein%20sequencing%20can%20be%20categorized.breaking%20down%20polypeptides%20into%20peptides.(link to source))

Evolution in protein sequences

Protein sequence evolution is an extraordinarily complex process, whose roles are ultimately dependent on the needs of the living organisms.

[https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006767\(link to source\)](https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006767(link to source))

Proteins that are more abundant are seen to evolve faster than proteins that are less abundant i.e increasing **expression levels** lead to reduced rates of protein evolution.

[https://www.pnas.org/doi/full/10.1073/pnas.1810194115#:~:text=Proteins%20that%20are%20more%20abundant,reduced%20rates%20of%20their%20evolution.\(link to source\)](https://www.pnas.org/doi/full/10.1073/pnas.1810194115#:~:text=Proteins%20that%20are%20more%20abundant,reduced%20rates%20of%20their%20evolution.(link to source))

The rate of protein sequence evolution (k) equals the rate of mutation (μ) multiplied by the proportion (p) of mutations that are neutral, as beneficial mutations are considered too rare to affect the rate of protein evolution the rate of protein sequence evolution (k) equals the rate of mutation (μ) multiplied by the proportion (p) of mutations that are neutral, as beneficial mutations are considered too rare to affect the rate of protein evolution. Therefore, $k = \mu * p$.

[https://www.nature.com/articles/nrg3950#:~:text=According%20to%20the%20well%2Daccepted,the%20rate%20of%20protein%20evolution.\(link to source\)](https://www.nature.com/articles/nrg3950#:~:text=According%20to%20the%20well%2Daccepted,the%20rate%20of%20protein%20evolution.(link to source))

The normal rate of mutation is 10^{-7} - 10^{-8} per nucleotide per cell division this estimate applies only to small-scale mutations, which produce mis-sense, nonsense or frameshift changes. <https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/mutation-rate#:~:text=The%20normal%20mutation%20rate%20is,sense%2C%20nonsense%20or%20frame%20shift%20changes>.(link to source)

A nucleotide is the basic building block of nucleic acids (RNA and DNA). Protein sequences diverge with time (this is what evolution means). Schematically, these changes may relate to drift or adaptation: Drift – Sequence changes occurring due to random sampling while preserving the protein's structure and function. Changes in protein properties including the acquisition of new biochemical activities.

[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9214755/#:~:text=Protein%20sequences%20diverge%20with%20time,function%20\(see%20purifying%20selection\)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9214755/#:~:text=Protein%20sequences%20diverge%20with%20time,function%20(see%20purifying%20selection)).(link to source).

Like most organisms, evolution of cells depends on the change in the requirements of the human body.

Protein structure

Introduction

The primary structure of a protein is the sequence of amino acids chained together to form polypeptide chain. Each amino acid is linked to the next amino acid through peptide bonds created during the protein biosynthesis process.

<https://www.ncbi.nlm.nih.gov/books/NBK564343/#:~:text=Fundamentals,during%20the%20protein%20biosynthesis%20process>.(link to source)

In order to understand protein structure, it is necessary to know that there are 4 levels of protein structure- primary, secondary, tertiary and quaternary.

Primary structures-The primary structure of a protein is its amino acid sequence; this sequence determines a protein's secondary structure. It dictates the 3-D conformation the folded protein will have.

Secondary structures- The two main types of secondary protein structures are the alpha helix and beta pleated sheet. The beta sheet (β -sheet, also β -pleated sheet) is a common motif of the regular protein secondary structure. Beta sheets consist of beta strands (β -strands) connected laterally by at least two or three backbone hydrogen bonds, forming a twisted, pleated sheet. [https://en.wikipedia.org/wiki/Beta_sheet#:~:text=The%20beta%20sheet%20\(%CE%B2%2Dsheet,a%20generally%20twisted%2C%20pleated%20sheet](https://en.wikipedia.org/wiki/Beta_sheet#:~:text=The%20beta%20sheet%20(%CE%B2%2Dsheet,a%20generally%20twisted%2C%20pleated%20sheet)).(link to the source)

An alpha helix (or α -helix) is a sequence of amino acids in a protein that are twisted into a coil (a helix).

The alpha helix is the most common structural arrangement in the secondary structure of proteins. https://en.wikipedia.org/wiki/Alpha_helix(link to source)

Tertiary structures-The tertiary structure of a protein refers to the overall three-dimensional arrangement of its polypeptide chain in space.

Quaternary structures-The quaternary structure of a protein is the association of several protein chains or subunits into a tightly packed arrangement.

<https://www.sciencedirect.com/topics/chemistry/quaternary-structure#:~:text=The%20quaternary%20structure%20of%20a%20protein%20is%20the%20association%20of,forces%20between%20nonpolar%20side%20chains.>(link to the source)

Evolution in protein structure

Protein structures evolve through various combinations of mechanisms which often include gene duplication before mutation and selection.

<https://www.sciencedirect.com/science/article/abs/pii/S09594440X0200338X>(link to the source)

Duplication is a type of mutation that involves the production of one or more copies of a gene or region of a chromosome. A protein can be unfolded, or denatured, by treatment with certain solvents, which disrupt the noncovalent interactions holding the folded chain together. A change in the gene's DNA sequence may lead to a change in the amino acid sequence of the protein. This treatment converts the protein into a flexible polypeptide chain that has lost its natural shape. Even changing just one amino acid in a protein's sequence can affect the protein's overall structure and function.

<https://www.ncbi.nlm.nih.gov/books/NBK26830/#:~:text=A%20protein%20can%20be%20unfolded,has%20lost%20its%20natural%20shape.>(link to the source)

Generally, the primary sequence of a protein dictates its fold and function, and cumulative changes in this sequence lead to the evolution of protein structures and functions.

<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/evolution-of-protein-structure#:~:text=Generally%2C%20the%20primary%20sequence%20of,sequences%20fold%20into%20different%20structures.>(link to source)

The Histone proteins are among the most conserved in all organisms. These proteins carry a very important function, which is the packing of DNA, and therefore they can withstand very few mutations in their evolutionary history. The speed at which a protein evolves depends on how stable a protein's folded structure is, how well it avoids aggregation, and how well-chaperoned it is.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6140491/#:~:text=The%20speed%20at%20which%20a,how%20well%2Dchaperoned%20it%20is.>(link to the source)

It is possible that the genomic mutations needed for conferring a novel function might not be present in a population, they can however, rise by non-genetic mechanisms mediated by errors in replication, transcription, and translation (phenotypic mutations).

<https://onlinelibrary.wiley.com/doi/full/10.1002/pro.4362>(link to the source)

Phenotypic mutations (errors occurring during protein synthesis) are orders of magnitude more frequent than genetic mutations. Consequently, the sequences of individual protein molecules transcribed and translated from the same gene can differ.

[https://www.pnas.org/doi/full/10.1073/pnas.0809506106#:~:text=Phenotypic%20mutations%20\(errors%20occurring%20during,the%20same%20gene%20can%20differ.](https://www.pnas.org/doi/full/10.1073/pnas.0809506106#:~:text=Phenotypic%20mutations%20(errors%20occurring%20during,the%20same%20gene%20can%20differ.)(link to the source)

Protein function

Introduction

The sequence of amino acids determines each protein's unique 3-dimensional structure and its specific function. It helps repair and build your body's tissues, allows metabolic reactions to take place and coordinates bodily functions. In addition to providing your body with a structural framework, proteins also maintain proper pH and fluid balance..

<https://www.healthline.com/nutrition/functions-of-protein#:~:text=Protein%20has%20many%20roles%20in,proper%20pH%20and%20fluid%20balance.>(link to the source)

There are a total of seven different protein types under which all proteins fall. These include antibodies, contractile proteins, enzymes, hormonal proteins, structural proteins, storage proteins, and transport proteins.

<https://www.physio-pedia.com/Proteins#:~:text=%3B%20Tryptophan%3B%20Valine.-.Types%20of%20Proteins.storage%20proteins%2C%20and%20transport%20proteins.>(link to the source)

Here is the explanation to them all:

- a) Antibodies- Antibodies are proteins that protect you when an unwanted substance enters your body.
- b) Contractile proteins- Contractile proteins are the proteins that control the sliding of muscles.
- c) Enzymes- Enzymes are the proteins that help in speeding up the chemical reaction of the body.
- d) Hormonal proteins- These proteins send out receptors to our brain initiating growth, development and other functions
- e) Structural proteins- Structural proteins are proteins that living organisms use to maintain their shape or structural integrity. Some common structural proteins are keratin, collagen, actin, and myosin.
[https://www.studysmarter.co.uk/explanations/biology/biological-structures/structural-proteins/#:~:text=Structural%20Proteins%20Definition.-Organic%20compounds%20are&text=Structural%20proteins%20are%20proteins%20that,or%20monomers%2C%20called%20amino%20acids.\(link to source\)](https://www.studysmarter.co.uk/explanations/biology/biological-structures/structural-proteins/#:~:text=Structural%20Proteins%20Definition.-Organic%20compounds%20are&text=Structural%20proteins%20are%20proteins%20that,or%20monomers%2C%20called%20amino%20acids.(link%20to%20source))
- f) Storage proteins- Storage proteins are functionally defined as proteins that are synthesized during feeding, when amino acid availability is in excess of demand, and depleted to provide substrates during periods of high demand, such as moulting.
- g) Transport proteins- These proteins are the proteins that transport substances like food around the body.

In short, proteins control each and every aspect of our existence and make up the very fundamentals of our existence.

Evolution in Protein functions

Even very tiny, subtle changes in protein structure can give rise to big changes in the behaviour of proteins. Haemoglobin, for example, undergoes an incredibly small structural change upon binding of one oxygen molecule, and that simple change causes the remainder of the protein to gain a considerably greater affinity for oxygen that the protein didn't have before the structural change.

[https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3ABiochemistry_Free_For_All_\(Ahern_Rajagopal_and_Tan\)/02%3A_Structure_and_Function/203%3A_Structure_Function-Proteins_I\(link_to_the_source\)](https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3ABiochemistry_Free_For_All_(Ahern_Rajagopal_and_Tan)/02%3A_Structure_and_Function/203%3A_Structure_Function-Proteins_I(link_to_the_source))

The cellular environment is a constantly evolving thing; proteins are required to quickly respond to changes in the protein environment. For this reason, cells have evolved various ways to modify or influence protein function.

[https://portlandpress.com/essaysbiochem/article/66/3/255/231650/Uncovering-protein-function-from-classification-to\(link_to_the_source\)](https://portlandpress.com/essaysbiochem/article/66/3/255/231650/Uncovering-protein-function-from-classification-to(link_to_the_source))

Many Changes in Proteins Are Driven by Phosphorylation

Enzymes are regulated by more than the binding of small molecules. A second method that is commonly used by eucaryotic cells to regulate a protein's function is the covalent addition of a phosphate group to one of its amino acid side chains.

[https://www.ncbi.nlm.nih.gov/books/NBK26911/#:~:text=Many%20Changes%20in%20Proteins%20Are,its%20amino%20acid%20side%20chains.\(link_to_the_source\)](https://www.ncbi.nlm.nih.gov/books/NBK26911/#:~:text=Many%20Changes%20in%20Proteins%20Are,its%20amino%20acid%20side%20chains.(link_to_the_source))

Proteins are built as chains of amino acids, which then fold into unique three-dimensional shapes. Bonding within protein molecules helps stabilize their structure, and the final folded forms of proteins are well-adapted for their functions.

[https://www.nature.com/scitable/topicpage/protein-structure-14122136/#:~:text=Proteins%20are%20built%20as%20chains,well%20adapted%20for%20their%20functions.\(link_to_the_source\)](https://www.nature.com/scitable/topicpage/protein-structure-14122136/#:~:text=Proteins%20are%20built%20as%20chains,well%20adapted%20for%20their%20functions.(link_to_the_source))

The amino acid sequence determines protein structure and structure dictates biochemical function, proteins that share a similar amino acid sequence usually perform similar biochemical functions, even when they are found in distantly related organisms.

[https://www.ncbi.nlm.nih.gov/books/NBK26820/#:~:text=Because%20amino%20acid%20sequence%20determines,found%20in%20distantly%20related%20organisms.\(link_to_the_source\)](https://www.ncbi.nlm.nih.gov/books/NBK26820/#:~:text=Because%20amino%20acid%20sequence%20determines,found%20in%20distantly%20related%20organisms.(link_to_the_source))

Battling Parkinson's Disease: Cause, Affection, Symptoms, Treatments

By Arya Devanath

Abstract

Parkinson's disease (PD) is a complicated condition characterized by both neurologic and nonmotor symptoms. PD affects 1% of the population and is caused by dopaminergic degeneration of the basal ganglia. Dopaminergic neurons of the substantia nigra degenerate, which is its defining characteristic. Abnormal balance, gait, and falls are common issues that might exist during diagnosis and worsen with time. This degeneration causes motor and non-motor symptoms, which lowers the quality of life for those affected. Globally, the prevalence of PD is rising, which presents a severe health risk. The motor characteristics of PD, such as bradykinesia (muscle stiffness) and resting tremor, are collectively referred to as parkinsonism. Although there are other secondary causes of parkinsonism, including disorders that resemble PD and drug-induced causes, dopamine neurotransmitter loss is the most prevalent cause of the condition.

Introduction

James Parkinson's essay on Shaky Palsy, published in 1817, contained the earliest clinical description of PD. Jean Martin Chorot subsequently changed the name "Paralysis Agitans" to PD. A groundbreaking discovery in 1960 that PD patients had lower striatal dopamine (DA) levels resulted in several amazing breakthroughs and expanded our knowledge of pathogenesis. This included discovering that Lewy bodies (abnormal protein clumps) increased, and pigmented neurons that produce DA decreased.

Even though PD has been reported to mainly affect people older than the age of 60, people in their 30s and 40s can also develop PD. Because of the neuroprotective effects of estrogen on the nigrostriatal dopaminergic system, the diagnosis of PD differs by gender at a ratio of 3:2 in favor of males.

The presence of nonmotor features suggests the involvement of other neurotransmitters (chemicals that allow neurons to communicate) of the glutamatergic, cholinergic, serotonergic, and adrenergic systems, in addition to dopaminergic systems. According to further data, PD may start in the anterior olfactory nucleus, the dorsal motor nucleus of the vagal and glossopharyngeal nerves, or the brain stem. This suggests that the illness might start in the brain stem and then move up to higher parts of the brain.

The absence of sensitive and accurate methods to track pathogenesis has been one of the biggest obstacles to creating valid neuroprotective medicines. Nowadays, clinical research is being done to test several immunotherapies, including immunization, protein clearance methods, and antibodies that prevent protein aggregation or target the protein synuclein. Patients with PD might feel hopeful about their future prognosis and results thanks to emerging medicines such as advanced surgical techniques, drug delivery methods, and medications that relieve symptoms.

Ninety genetic risk variations account for 16–36% of the heritable risk of PD that is not caused by a single gene. Meanwhile, 3–5% of PD in most groups is explained by genetic factors connected to known PD genes, called monogenic PD.

The goal of current therapy is to control symptoms. Research indicates that a multidisciplinary care team of movement experts, social workers, pharmacists, and other medical professionals may benefit PD patients. Individual differences in treatment objectives highlight the importance of personalized management. Although there is currently no treatment to stop or slow down the course of PD, several intriguing approaches are being investigated for their ability to change the condition in light of recent discoveries about the genetic origins and processes of neuronal death.

Causes of PD

Although the exact etiology of PD is still unknown, experts think that a person's lifestyle choices, genetic makeup, and environmental circumstances affect how the disease develops. Two unchanging features—the increased presence of Lewy bodies (LB) and a decreased number of dopaminergic neurons in the substantia nigra pars compacta (SNc)—must be pathologically confirmed for the diagnosis of PD. A mere 1% of PD individuals have types of Mendelian pattern, in which they have an apparent genetic pattern that runs in families. However, 15% of cases have loose family aggregation, where more than one person can be affected. Genetic mutations contribute 10–15% to PD, whereas environmental hazards, including head trauma, residential location, pesticide exposure, MPTP exposure, etc, cause 90–95%.

Age

Growing older, typically people about 60, is one of the most prevalent risk factors because aging alters cellular processes that increase the risk of neurodegeneration related to PD pathogenesis. PD progresses more quickly when compensatory mechanisms fail and age-related somatic damage occurs. Buildups of aberrant and misfolded alpha-synuclein (α -synuclein), a neuronal protein, are linked to every instance of PD. This misfolded protein accumulates and clusters into collections, which then come together to create microscopic protein threads known as fibrils. These are the components that make LBs. PD individuals with LBs have clumped versions of α -synuclein in their spinal fluid that are indestructible by cells. Extended infusions of rotenone (chemical compound) lead to symptoms resembling PD, including the specific nigrostriatal pathway dopaminergic neuron depletion and the development of abnormal structures similar to LBs, which contain α -synuclein.

Due to changes in medication effects, an increased risk of dementia, and nursing home admissions, aging has an impact on the clinical presentation of PD. In mild cases, patients typically have reduced function due to age-related dopaminergic activity, incidental LB disease, and degenerative and vascular diseases. Patients' brain areas become more vulnerable as they age, disintegrating regular cellular repair and compensation systems. Since aging is unpredictable, unrepaired cellular damage may accumulate due to predictable and unpredictable

factors. The aging process involves interactions between genes, environment, and cellular and molecular changes that decide which cells age, are fortunate, or die of neurodegeneration. Furthermore, stress, poor food, dehydration, and inactivity may intensify PD.

MPTP

MPTP-MPP⁺ (an organic compound) harms the SNc when converted by the brain. Contact with MPP⁺ results in severe, irreversible parkinsonism. Exposure to metal manganese and certain herbicides, such as paraquat and maneb, enhances PD risks. Derivatives of tetrahydroisoquinoline and B-carboline, which are structurally linked to MPTP and can be found in some foods, have been identified in PD patients' brains and cerebrospinal fluids. These compounds cause nigrostriatal damage.

DA depletion

The deficiency of striatal DA resulting from the selective degeneration of nigrostriatal neurons is the primary cause of the motor impairments associated with PD. Greater levels of neuronal discharge in the globus pallidus internus (Gpi) and subthalamic nucleus (STN), as well as slightly reduced discharge in the globus pallidus externus (Gpe), are linked to the depletion of striatal DA in human PD and animals models of MPTP-induced parkinsonism. Developments in several fields have lately provided positive findings relating to the pathophysiology of motor dysfunction in PD. Depletion of striatal DA in PD affects the intricate networks of cortical, basal ganglia, and thalamic neurons. These networks comprise numerous layers and the borders of looping structures, many neuronal elements, various neurotransmitters, neuromodulators, axonal and dendritic branching patterns, and layer-to-layer connectivity.

"Contiguous" but "functionally" specialized pathways form groups of circuits or loops that pass through corresponding basal ganglia portions and are redirected to specific cortical fields. These circuits connect the basal ganglia, the thalamus, and the cerebral cortex, which emerge from a group of cortical areas connected physically and functionally. Each loop's information processes are serviceably divided, supporting an independent array of cognitive functions. Thus, skeletal and ocular motions are controlled by sensorimotor and oculomotor networks. Behavior set-switching and executive function is controlled via the ventromedial and dorsolateral prefrontal circuitry. And the limbic circuit controls the processing of emotions.

Several studies suggest that some functions may go astray in PD due to decreased information processing resulting from striatal dopaminergic depletion. Like this, hypokinesia and parkinsonian stiffness in PD are caused by insufficient inhibition of thalamic targets due to disinhibition of GPi and SNr. This is caused by decreased dopaminergic activity in D1 receptors and enhanced glutaminergic inputs inhibiting GPe. Increased glutaminergic drive from STN raises GPi or STN neuron activity, lowering thalamocortical activity. Instead of concentrating on status variables like mean firing rates, contemporary methods are centered around dynamic elements of shifts in neuronal activity in PD, such as oscillatory busting and synchronization of discharge among nearby neurons. Recordings from Parkinsonian monkeys and PD patients show

rhythmic neuronal bursting in GPe and STN, with low-frequency oscillatory field potentials. Normally, GABAergic feedback restrains oscillatory interactions between GPe and STN, but weakened feedback increases synchronization and oscillatory activity associated with motor and cognitive impairments in PD.

The information that passes through the basal ganglia, essential for movement, is influenced by DA signals from the brainstem. DA receptors are members of the d1 and d2 subclasses, part of a group of essential proteins in cellular communication. These are determined by their effect on the cytoplasmic cyclic adenosine monophosphate (a molecule critical for cellular signaling), either increasing (d1) or decreasing (d2). These effects are essential for the movement-related information of the basal ganglia processes. A protein called DARPP-32, modulated by intracellular signaling pathways, controls the activity of glutamate and GABA receptors. DA plays a role in this process by affecting the phosphorylation (a chemical modification) of DARPP-32. It strongly and partially impacts the speed at which the fast-acting glutamate and GABA coincide synaptically. This control is implemented via DARPP-32's modulation of the phosphorylation of synaptic receptors associated with those neurotransmitters. D1 receptor activation raises phosphorylation, increasing the sensitivity of GABAA, AMPA, and NMDA glutamate receptors. However, the phosphorylation and sensitivity of those receptors to their corresponding agonists (specific messengers) are reduced upon activation of d2 receptors. When glutamate interacts with AMPA receptors, the postsynaptic (cell membrane) depolarizes and removes a magnesium block of NMDA receptors. This allows many sodium ions to enter the cell and only a few potassium ions to exit the neuron. This disrupts the neuronal excitability and firing patterns, reducing DA release.

Additionally, movement disorders are caused by extracellular zinc (Zn^{2+}), which significantly affects nigral dopaminergic neurons because of AMPA receptor activation. This suggests that AMPA receptors that allow calcium (Ca^{2+}) and zinc (Zn^{2+}) to pass through and that lack a component called GluR2 play a small role in the management of PD. Motor problems in PD patients have been caused by levodopa increasing the number of glutamate receptors in the medial putamen (brain region). So, AMPA receptor antagonists (drugs that block AMPA receptors) might help reduce PD symptoms.

Ubiquitin system

Moreover, by getting rid of abnormal proteins, the ubiquitin system is essential for preserving cellular health. The chemical ubiquitin identifies specific proteins in cells for destruction by cell structures called proteasomes, which initiate chemical events that break down those proteins into smaller parts. Because toxins and other substances can build up to dangerous amounts that trigger cell death, the disposal system is unable to function. Impairment in the ubiquitin-proteasome system (UPS) plays a significant role in neurodegenerative disorders. UPS helps decompose misfolded α -synuclein and regulates the communication between various genes. α -synuclein prevents the UPS from operating correctly, which leads to an accumulation of α -synuclein and further reduction of UPS activity.

Stress

According to pathological research, inflammation, proteolytic, and oxidative stress are significant factors in PD pathogenesis. People with PD have altered brains due to oxidative stress caused by free radicals damaging DNA, proteins, mitochondria, and lipids, which are vital. Rising oxidative stress levels promote the cytoplasmic release of proapoptotic cytochrome c (a form of protein) from damaged mitochondria. This triggers apoptosis, a programmed cell death (PCD) within the mitochondria. This cell death is related to the death of neurons during PD. Specific characteristics of DA neurons in the midbrain can exacerbate oxidative damage in SNc. Because of the increased iron levels within the SNc, DA outside the cells can induce oxidative stress inside nigral neurons through several mechanisms. This progressive intracellular iron buildup develops in PD in SNc neurons and microglia, causing harmful reactions and damage. When this accumulation of iron occurs, ferritin levels are raised concurrently. Ferritin helps to balance off the potential danger of localized redox toxicity associated with higher iron levels. However, the anticipated rise in local ferritin does not happen in PD. Research indicates that DA in the ventral tier of SNc have lower concentrations of neuromelanin (NM) than their more resistant dorsal tier counterparts. NM can attract iron and bind to both the active Fe^{2+} (iron) and inactive Fe^{3+} , reducing the possibility of oxidative stress by oxidizing or chelating.

Proteolytic stress results from accumulating an undesired protein. Proteolytic stress-related protein aggregation may exacerbate or trigger PD neurodegeneration. Subsequently, practically all of the α -synuclein in the presynaptic terminal (the end of a neuron's axon) is still unfolded, with the remaining portion being concentrated near the synaptic vesicles (a structure that stores neurotransmitters). When they attach to the plasma membrane, these generally take the form of an α -helix (spiral), raising the possibility that α -synuclein contributes to synaptic vesicle renewal and preservation. However, an increase in α -synuclein quantities exerts a greater strain on the UPS to remove and break down proteins, intensifying proteolytic stress overall.

Furthermore, the natural form of α -synuclein binds to the critical anti-apoptotic protein 14-3-3, which weakens its defenses against the activation of PCD. In its advanced form, unfolded α -synuclein creates cytotoxic protofibrils called B-pleated sheets. Protofibrils could enhance synaptic vesicle permeability (allowing substances to pass through), leading to DA leaking into the cytoplasm and increasing oxidative stress. Because more oxidized and nitrated proteins need to be removed by UPS, oxidative stress may cause proteolytic stress to worsen.

Microglia

In the regions where neuron loss in PD and MPTP-induced parkinsonism are present, noticeable inflammation often occurs. Reactive astrocytes and activated microglia (support cells) comprise most inflammatory cells. In PD, microglia are also thought to aid neurodegenerative processes. Locations where neurons die and are lost experience a buildup and firing of microglia. Activated microglia is the primary cause of the increase in inducible nitric oxide synthase (iNOS) levels in the Parkinsonian nigra. Long-term production of nitric oxide (NO) is linked to

iNOS activation. The NO molecules can easily permeate cell membranes and reach SNc neurons. Once it interacts with locally generated superoxide anions (harmful molecules), its reaction forms peroxynitrite (highly reactive harmful molecules). This may worsen the harm that oxidative stress causes to nigral neurons' internal fluid, proteins, and DNA. Cytokines, released by activated microglia, can intensify inflammatory reactions and stimulate other microglia nearby. Many of these molecules, including tumor necrosis factor- α (TNF- α), have been found in the nigral tissue of PD patients. Microglia-derived TNF- α can activate a part of the TNF receptor called the "death domain" by binding to TNF receptor 1 on the surface of SNc neurons. This initiates the extrinsic PCD pathway, leading to cell death through initiator caspase 8 and executioner caspases (proteins involved in PCD).

Lewy Bodies and α -synuclein

The areas of the medulla oblongata, pontine tegmentum, and olfactory bulb are where LBs are first seen in the brain. The postmortem pathology of PD may be classified into six levels using a 6-tier system. The substantia nigral regions, basal forebrain, and neocortex are where LBs spread during the Braak stages 3 and 4, whereas stages 1 and 2 are presymptomatic and involve fewer LB-affected regions. The idea that PD begins in the autonomic nervous system is supported by the finding that LBs and α -synuclein aggregates are present in the salivary glands of the gastrointestinal tract nervous system before alterations in the brain. Researchers at the Udall Center at the Perelman School of Medicine of the University of Pennsylvania injected mice with synthetic forms of abnormal α -synuclein, eventually discovering that the brain was affected by misfolded α -synuclein. The researchers hypothesized that PD-like injection triggers the mouse's proteins to misfold, transmitting symptoms from cell to cell. Regarding age variables, LB development serves as a signal for defense mechanisms against age-related neural system deterioration and dysfunction.

Genes

Several genes, such as SNCA, PARK 2, SPARK 6, PARK 5, PARK 7, PINK 1, LRRK 2, GBA, and NR4A2, contribute to the development of PD. Early PD is affected by SNCA mutations. These mutations assemble LB clusters and produce the protein α -synuclein. The α -synuclein gene is thought to misfold, producing too many protein copies.

PARK 2 aids in the recycling and decomposition of cells by producing the parkin protein. Its slower disease development makes it the most prevalent gene for early-onset PD in those 50 or younger. SPARK 6 is similar but more rare. PD is linked to both mutations in an autosomal recessive manner.

PARK 7 produces the protein DJ-1, which fights oxidative and mitochondrial stress. This stress causes unstable chemicals to build up to the point where they may eradicate cells. The DJ-1 gene strengthens the cell's defenses against metal toxicity. While mutations in PARK 5 can raise the risk of late-onset PD, mutations in PARK 7 can produce a rate of early-onset PD.

When interacting with parkin, PINK 1, a protein produced by the gene PTEN-induced putative kinase 1, helps protect the mitochondria from stress (mitophagy). Together, they can detect mitochondrial damage and attract parkin, which initiates mitophagy. To aid in developing medicines, researchers are figuring out how to activate the PINK 1/parkin pathways to promote mitophagy, where mutations arise in early PD. Parkin contributes to several other mechanisms that lead to PD, including sporadic variants that are toxic to α -synuclein. There may be a connection between parkin and sporadic PD since data suggests that parkin becomes inactive in sporadic types of PD.

Leucine-rich repeat kinase 2 (LRRK 2) is also a protein kinase; however, mutations in this gene have been connected to late-onset PD. This is the most prevalent gene, occurring in 10% of hereditary forms and 4% of cases of no family history. It affects the production and elimination of undesirable proteins and is present in early and late-onset PD types. Kinase controls protein enzymes by phosphorylation. Its mutations enhance the pace at which the ribosomal proteins—a crucial part of the machinery that makes proteins—are tagged by the gene's protein, allowing the machinery to produce excessive proteins and ultimately causing cell death. It is thought to prevent autophagy, where nutrients are broken down, cellular components are recycled, essential quality control mechanisms are removed, and aberrant proteins and damaged, useless waste organelles are eventually eliminated by the cells. This kind of autophagy is called chaperone-mediated autophagy, in which a damaged protein is escorted to lysosomes, which contain acids that aid in decomposing undesirable molecules.

The beta-glucocerebrosidase gene is GBA. Mutations in the lysosomal enzyme-encoding gene, GBA, are linked to Gaucher's disease, a lysosomal storage illness. Parkinsonism is more common in those who have this illness than in those who do not. According to NIH-funded research, there is a five-fold increased risk of GBA mutations in PD patients compared to individuals without confirming the connection between Gaucher's and PD. Likewise, neurodegeneration and α -synuclein buildup result from GBA depletion.

The protein that NR4A2 encodes controls the transcription of the TH gene, which is essential for the formation and growth of dopaminergic neurons in the midbrain. This gives importance to inheritance patterns that vary based on the underlying genes. For instance, autosomal dominant pathology occurs when genes inherited from a single parent, such as SNCA or LRRK2, need one mutated gene copy to cause the illness. However, when two copies of the gene in each cell are mutated for the condition to manifest, it is referred to as the autosomal recessive pattern for PARK 2, PARK 7, and PARK 1. These genes alter the PD probability in sporadic cases.

Down Syndrome

It has also been established that Down syndrome (DS) contributes to the development of PD. 23 DS patients, ages 49 to 74, participated in a clinicopathological investigation to determine the relationship between DS and PD. It was discovered that two individuals, ages 50 and 56, had cortical Alzheimer-like disease with LB development in the SNc. 20% of the 53

participants in another study who had dementia and DS also reported extrapyramidal symptoms (movement control). Two PD patients who showed indicators of dementia 1.5 years later had tissue loss in the temporal lobe. 36% of patients had extraperitoneal characteristics, such as DS and Alzheimer's disease (AD), and half of the patients responded to therapy.

Nonetheless, a 45-year-old man with levodopa-responsive DS and PD had a loss of memory, initiative, and focus. For the past two years, a young woman with DS displayed characteristic Parkinsonian symptoms and continued to respond to levodopa. This patient is special since her parkinsonism started at a young age; she responded dramatically and persistently to levodopa with no cognitive deterioration.

Extrapyramidal symptoms in DS patients may be misdiagnosed with psychomotor symptoms (mental or motor issues) caused by AD, which might lead to an inadequate reporting of Parkinsonism cases. However, discipline-based methods help identify related conditions. For instance, if a neurological examination reveals Parkinsonian signs in a DS patient, levodopa is advised as a recommended trial.

Alcohol

Alcohol can exacerbate bradykinesia, stiffness, and tremors and alter motor function and coordination by interacting with PD symptoms. Since it depresses the CNS, it can worsen cognitive impairment and raise the chance of falling. It can interfere with PD treatments, lessen their efficacy, and aid in the disease's quicker development. Alcohol interferes with sleep cycles while affecting carbidopa-levodopa, too. However, because alcohol contains protective properties, it can be consumed in moderation, depending on the variety.

REM sleep

Another feature of PD pathogenesis is the cholinergic neuron depletion in the pedunculopontine nucleus (PPN), which includes cholinergic neurons in the pars compacta and glutamatergic neurons in the pars dissipatus. While glutamatergic neurons project to the GPi, substantia nigra pars reticulata, SNc, and STN, the cholinergic neurons project to the thalamus and GPi/SNr. REM sleep is facilitated by cholinergic neurons found in PPN and laterodorsal tegmental nuclei, which receive converging monoaminergic inputs from the NMR, LC, and SNc neurons. The loss of those neurons increases REM sleep induction. Regular regulation of REM sleep with atonia (absence of muscle tone in REM) may depend on GABAergic input to the PPN from the basal ganglia. Atonia triggered by PPN activation can be suppressed during REM sleep by electrical microstimulation of SNr. PD's frequent occurrence of REM sleep behavior disorder may be explained by the degenerative loss of PPN neurons, which includes excessive and oscillatory GABA(a) mediated inhibition via SNr afferents. So, akinesia in PD is caused by the excessive GABAergic inhibitory outflow from GPi or SNr to PPN.

Synaptic Plasticity

Musculoskeletal pain is a common PD symptom, as are articular or arthritic pain, neuropathic pain, and radicular pain. Synaptic plasticity, such as long-term potential and depression (LTP/LTD), has been claimed to contribute to the pathophysiology of PD partly. The reduced activity of Casp3 (a protein that helps in PCD) prevents the formation of synaptic plasticity, which is required to induce LTD. This synaptic plasticity is essential for the induction of LTD, and when it is downregulated, the mechanism of chronic pain is disrupted. Alterations in pain-related cortical areas are thought to be brought on by neurotransmitter DA and serotonin imbalance, which regulate the brain's ability to process information. As the disease progresses, the oscillations become more pronounced, leading to more pronounced symptoms. Instead, DA may act directly on synaptic transmission mechanisms, such as through the regulation of calcium channels, or on other neurotransmitters, such as glutamate or GABA, which regulate LTP. The regulation of DA is essential for excitatory and inhibitory signals in the brain. D1-like receptors help regulate glutamatergic excitatory and GABAergic inhibitory signals onto pyramidal cells of the anterior cingulate cortex (ACC). Proper functioning of these receptors is essential for the brain's normal functioning, so abnormal functioning can lead to the pathophysiology of PD.

Other causes

An RNA-binding protein called FMRP regulates translational efficiency, which leads to synaptic flexibility. FMRP is a crucial messenger in the intracellular DA signaling pathways that modulate DA in the prefrontal regions. FMRP plays a role in DA modulation by regulating the surface expression and phosphorylation of DA receptors and DA transporters. Furthermore, the loss of FMRP leads to deficits in DA-related behaviors such as reward and reinforcement, suggesting that it is essential for DA-related behaviors.

The loss of nMR neurons that produce serotonin (5-HT) in PD leads to serotonin shortages in the cerebral cortex, brainstem, spinal cord (SC), and basal ganglia, which is associated with clinical depression.

Effect on the Brain

Because of the complex connections between the striatum and the prefrontal cortex, people with PD frequently show difficulties in activities requiring this part of the brain. Neuron loss in PD progressively affects the temporal and frontal lobes, compensating for movement, personality, behavior, and language abnormalities. The deterioration in cognitive performance can be attributed to the loss of tissue in these lobes. Although PD has historically been thought of as a brain disorder caused by α -synucleinopathy, data from experimental models using autopsy human and animal models suggests that the SC is impacted.

Substantia Nigra

The SNc is the brain area most impacted by the loss of DA neurons. The SNc is the primary location for synthesizing DA neurotransmitters and is also in charge of transmitting signals that direct and coordinate movement. DA is produced by melanin-containing neuron

clusters in the SNc. The term "basal ganglia" describes the deep-seated nuclei in the hemispheres of the brain, such as the globus pallidus, caudate-putamen, and striatum, which contain a reliable marker of dopaminergic neurons. Other susceptible cell populations in the brainstem are affected by related pathology outside of the SNc by the time the SNc becomes associated with PD. These include the median raphe's serotonergic nuclei and the locus coeruleus's noradrenergic neurons (LC). Except for the basal ganglia, the whole CNS receives nerve connections by LC neurons via noradrenergic fibers (regulates physiological functions). When LC neurons in the cerebellum and frontal cortex are lost, norepinephrine (neurotransmitter) levels in these regions are significantly reduced. The subcortical nuclei, which primarily aid in motor control, lose their DA connections due to the loss of DA-producing neurons. The degeneration of SNc's dopaminergic neurons also sets off a series of functional alterations that impact the whole ganglia network. This clarifies why a person experiencing basal ganglia dysfunction may struggle to initiate, halt, or maintain movement. As a final common mechanism of neuron death, the three pathways leading to the activation of PCD are also thought to be involved, including oxidative stress, proteolytic stress, and local inflammation.

The midbrain is crucially connected to the forebrain and hindbrain, composed of pairs of massive nerve fiber bundles, with the SNc lying between the bundles and the tegmentum (a part of the midbrain). LBs are detected in the SNc neurons of PD when 60–80% of the SNc neurons are damaged. Other extra nigral neuron losses include acetylcholine, serotonin, noradrenaline, and glutamate (neurotransmitters). The brain's systems that control the chemical pathways dependent on acetylcholine, serotonin, and norepinephrine are damaged. The posterior neostriatum (the sensory portion of the putamen) receives dopaminergic innervation from a subpopulation of nigrostriatal neurons. The sharpest decrease in striatal DA levels was also impacted by neuronal death, which explains the pattern of DA depletion.

Furthermore, while the remainder of the nucleus is mostly intact and the A8 group of DA neurons is entirely left alone, there is a substantial decrease in neurons in the ventrolateral SNc. In mesocortical areas, dopaminergic depletion in the cortical layer occurs, but there are significant amounts of dopaminergic remnants in the deep layer of the cortex. This depletion results in pattern-evoked potentials and abnormalities in electroretinography, indicating disruptions in DA regulation.

Emotional brain regions

The emotional components of the human body are documented as symptoms of PD, which affects brain regions that include the periaqueductal gray, striatum, insula, and amygdala. The primary emotion of the amygdala is fear, which links emotions to sympathy, violence, and addiction-related behaviors. Twenty percent of its volume is lost during PD, which results in severe pathological changes. As emotions are affected, mood swings lower the quality of life.

According to studies, the basolateral nucleus, a region in the amygdala, could be in charge of producing PD-related visual hallucinations. This may be due to its function in controlling DA levels, which are linked to regulating visual hallucinations. Since LBs are linked

to visual hallucinations, the basolateral nucleus might also cause a higher percentage of LB-containing neurons. Two times as many basolateral neurons are present in PD patients who have visual hallucinations. In PD, the basolateral amygdala nucleus is frequently pathologically involved. This nucleus' glutamatergic neurons are connected to the ventral striatum and ventromedial prefrontal cortex, two regions important for memory and learning. Neuron loss due to LB dysfunction impacts the amygdala's cortical nucleus—which has a strong bilateral relationship with olfactory structures—and contributes to hyposmia (reduced sense of smell) in PD. The number of DA neurons in the olfactory bulb rises. In 90% of PD patients whose olfaction is officially examined, hyposmia is evident and typically symptom-less. DA inhibits olfactory transmission at the glomerular level (refers to the glomeruli). Early and frequent olfactory impairment in PD is linked to significant LB pathology in the olfactory bulb, early neuron loss, and LB development in the anterior olfactory nucleus.

The striatum is a network of interconnected basal ganglia nuclei whose activities are associated with motor control. Within the dorsal striatum, many varieties of spiny projection neurons (SPNs) receive axons from the midbrain and DA-producing neurons from the SNc. These neurons are involved in motor movement programming and cognitive behavior programming. Dorsal striatal DA facilitates the establishment of a "response set" in motor preparation processes, leading to efficient responses and coordinating sensorimotor behavior involved in consummatory actions. DA in the ventral striatum facilitates reward-related cues in appetitive components of behavior.

Similarly, the GPe and the STN participate in the "direct" and "indirect" pathways that arrange the information flow through the basal ganglia. The "direct" route has a chemical called P/dynorphin that expresses dopamine D1-receptors. The "indirect" route includes enkephalin, which expresses D2-receptors. The dopaminergic output from the striatum to the SNc degenerates in PD. The reduction in striatal output via the "direct" pathway is accompanied by an increase in the activity of "indirect" striatal neurons and a decrease in striatal DA. These modifications enhance the GABAergic basal ganglia outflow to the thalamus.

Additionally, DA influences the long-term synaptic plasticity between glutamatergic inputs and medium spiny neurons in the striatum. As a result, degeneration of the nigrostriatal system in PD directly affects the activation level of medium-spiny neurons, which significantly alters the structure and function of synapses in the striatum. According to recent findings in the MPTP-treated nonhuman primate model of PD, the sensorimotor striatum was the most severely affected striatal region, displaying as much as 50% spine loss in cases with severe DA denervation. However, postmortem findings from human Parkinsonian brains have shown less damage to the ventral striatum. Fundamentally, there is a strong correlation between the extent of striatal DA degradation and the striatal spine loss in parkinsonism.

There is also evidence of the emergence and development of LBs in the tegmentum's periaqueductal gray (PAG) region. Lewy neurites (LN) and LBs heavily accumulate in the PAG as PD advances, making it more vulnerable to diminished activity and cognitive decline. The PAG is essential for transmitting and regulating pain, sympathetic reactions, and the acquisition

and execution of unpleasant and defensive behaviors. The sole nucleus, namely the ventrolateral tier, has been shown to have disrupted dopaminergic antinociceptive circuitry (neural pathways) in PD. Numerous investigations have demonstrated that central pain is caused by PAG damage. Both the facilitation and suppression of pain via specific pathways are included in chronic pain, and both can result in other illness states, including anxiety and depression. PAG's regions help modulate pain signals transmitted to the brain by the interactions between the neurotransmitters and receptors in signal transmission through the PAG. Those neurotransmitters include catecholamines, which help raise alertness in reaction to pain; serotonin and enkephalin, which help decrease the amount and intensity of pain signals; and GABAergic neurons, which help prevent the transmission of pain signals. Serotonin acts on 5-HT receptors to influence neurons in the PAG-rostral ventromedial medulla pathway. This pathway, in turn, influences the activity of enkephalin neurons, reducing pain perception by obstructing the brain's ability to receive pain signals. Significant LB and LN deposition in the PAG can cause functional impairments and lower pain thresholds due to decreased activity in PD. In the early stages of PD, the ventrolateral tier of the PAG is the only known nucleus with dopaminergic antinociceptive circuitry. This circuitry helps modulate pain by influencing descending projections that indirectly affect nociceptive impulses ascending to the dorsal horn of the SC. Reduction in ventrolateral periaqueductal gray activity was reduced by 50%, enhancing sensitivity to pain stimuli in PD.

Other connections

In PD, there is a notable decrease in the vital link between cognition and the connectivity involving the ACC. Cognitive impairment in PD is believed to be related to the functional connectivity between the ACC and the default mode network (DMN), a brain network active during restful states and self-referential thinking. The underlying cause of PD's impaired cognition is altered interactions between dopamine-related pathways (DAI), DMN, and the frontoparietal network (FPN). Test results related to the DAI's functional connection with the DMN indicate that cognitive performance was much worse in PD patients than in controls. The functional connection of the DAI with the ACC was dramatically diminished in PD, while in controls, cognitive performance was favorably correlated. The reduced connection between DAI and the ACC, coupled with the greater centrality of DMN and FPN in PD, was shown to be linked with cognitive impairment in PD.

Symptoms

The four primary symptoms that elderly people should look out for are tremors in hands, arms, legs, jaw, or the head; muscle stiffening, leaving it to remain contracted for an extended period; slowness of motion; and impaired coordination and balance that results in occasional falling. Other recorded symptoms include memory loss, pain or numbness, behavioral and emotional changes, difficulty in thinking, reasoning, swallowing, chewing, or speaking, frequent seizures, urinary problems, increased fear of falling, loss of constant sensitivity, neck tightness, small handwriting or unintentional writing, weight loss, sudden vision loss, skin problems,

slurred speech, and confusion. These behavior abnormalities and autonomic dysfunctions are caused by the lesional patterns in PD, which damage specific nuclear gray matter and its connections. Cognitive dysfunction can lead to amnesia, confusion in the evening hours, dementia, difficulty thinking and understanding, and an increase in anxiety and apathy. Nasal problems involve a distorted sense of smell, loss of smell, and facial stiffness.

When PD affects nerves outside the brain before moving to the brain, it manifests as constipation. There is a link between gastrointestinal problems and PD. The enteric nervous system governs the digestive system by regulating gastric acid production, which dictates the flow of the gastrointestinal (GI) tract. It has between 200 and 600 million neurons, ultimately separated into hundreds of myenteric and submucosal ganglia. Although symptoms can impact the GI system, the most frequent symptom is colon congestion.

One of the most prevalent symptoms is tremors, which manifest as a patient shaking rhythmically (a pill-rolling tremor), rubbing their thumbs and forefingers back and forth, or trembling while at rest but moving less when doing an action. Similarly, bradykinesia is a delayed movement condition that can make ordinary actions more difficult. Due to the freezing of the gait, patients would begin to walk with shorter steps and a propensity to drag or shuffle their feet. Furthermore, the abrupt reductions in blood pressure might result in neurogenic orthostatic hypotension, which can induce fainting. Depletion of DA causes a reduction in the capacity to conduct unconscious motions like smiling or blinking, ultimately leading to a lack of automatic movement. This causes patients to wake up sooner, disrupting their sleep habits and is the primary cause of nightmares, restless nights, and sleep disruptions.

Structural issues of the SC during PD include spinal stenosis, which causes pain when standing or walking and is caused by the shortening of the central spinal canal, which the SC passes through. The discomfort along the leg caused by radiculopathy (pinched nerve or sciatica) is brought on when the spinal nerves' exit pores become narrower. Other SC difficulties include bone spurs (bony protrusion), compression fractures (vertical collapse of the vertebral body), herniated discs (where the nucleus pulposus can push through the outer core and compress the holes), and arthritis (joint swelling). Another typical SC-impacted variable is lower back pain (LBP), where 43–74% of individuals with PD dystonia report aberrant muscle spasms brought on by PD. LBP can be exacerbated by aging, sadness, stiffness, and stooped posture, which worsen as activity levels decline. There seems to be a shared pathogenic link between the development of PD and lumbar spinal canal stenosis.

Everybody experiences distinct symptoms, some of which can worsen and change with time. It is essential to consider that drug interactions may potentially influence cognitive alterations. Hyposmia, sadness, anxiety, and sleep disturbances are examples of non-motor symptoms that arise in the preclinical stage prior to the manifestation of motor symptoms. Medication will become increasingly important or have unbearable adverse effects as it gradually advances. Levodopa, for instance, may thereafter result in dyskinesia and dystonia.

Additionally, individuals with Parkinson's disease dementia (PDD) who are diagnosed with PD often live for 5–7 years. PDD occurs when an individual with PD has severe cognitive

deterioration following a year or more of motor symptoms. In addition to increasing agitation, anger, and hostility, patients with associated dementia symptoms also start to experience visual perception issues. These consist of hallucinations, illusions, delusions, and misperceptions. This exacerbated their delusional and bizarre ideas, suspicious or mistrustful thoughts, and hearing, seeing, or sensing unreal things. Muscular degeneration goes hand in hand with difficulty locating or navigating objects correlating with the rise. Comprehending complex language and naming items will also be challenging for someone with PDD, further experiencing speech issues, such as difficulty forming words even when they know what they want to express.

Treatments

Although there is no known cure for PD, specific symptoms can be alleviated by medication, surgery, and other therapy. Medications can help regulate non-movement symptoms and raise DA levels in the brain by impacting neurotransmitters. Both functional neurosurgery and dopaminergic drugs can be used to treat PD symptomatically.

Levodopa

Levodopa is a first-line medication for the treatment of PD. Carbidopa, a DA booster, is typically given with levodopa to prevent and minimize its adverse effects. Levodopa dosage required to improve symptoms may be lowered as a result of these adverse effects, which can include nausea, vomiting, low blood pressure, and restlessness. Neurons use this to produce DA and refill the brain's depleting supply. Therefore, the cornerstone of PD treatment in rural areas is the continuous administration of levodopa. However, inconsistent usage may result in breathing difficulties, movement impairments, and the emergence of dyskinesias. Researchers postulate that levodopa may be linked to changes in the brain's blood vessels that modify a medication's capacity to cross the blood-brain barrier. Early diagnosis offers a treatment period in which motor deficits can be delayed or prevented before PD progresses. However, by diagnosis, 60–80% of SNc cells were no longer producing DA. Nevertheless, the yearly cost of treating PD is \$14 billion, and the additional expense of lost productivity amounts to \$6.3 billion every year. By 2040, the number of diagnoses is anticipated to have doubled.

The "OFF" period is when levodopa is not working as effectively, and typical PD symptoms reappear. This can result in dystonic discomfort, which might vary depending on when medications are taken. Usually, this may be resolved by adjusting and using various medicine doses. However, non-dystonic pain may exhibit "non-motor OFF" patterns (being an OFF phenomenon) that improve with levodopa dosage.

When taking PD medicines, people often experience changes in non-motor symptoms, like chronic pain, rather than motor symptoms. This type of pain can be managed by reducing the amount of time that a drug is off. Depression can exacerbate movement symptoms and contribute to chronic pain, as it is a typical non-motor symptom of PD. Specific therapies can reduce some pain, and when attempting to treat chronic pain in PD, a depression screening is typically

necessary. Discussing a referral to a psychiatrist about a neurologist is advised, as their plans may involve more than just medicine and psychotherapy.

Deep Brain Stimulation

For individuals who do not react well to therapy, deep brain stimulation (DBS) is an option, as it has been discovered to be one of the most successful drugs. However, despite DBS therapy, symptoms may still exist. During this surgical procedure, electrodes are implanted into specific brain regions—usually the globus pallidus or the STN—and attached to tiny electrical devices in the chest. An implanted pulse generator beneath the patient's collarbone through cables under the skin provides precisely regulated electrical signals to the DBS electrodes. It blocks messages to help halt PD symptoms associated with movement by gently stimulating certain regions that govern movement. Exercise, physical therapy (PT), and acupuncture are advised for LBP because they help determine the patient's optimal activity. It involves manual manipulation, pushing on muscle fibers, and moving joints. The PT evaluations look at movement patterns and their relationship to pain to address issues with poor posture or gait patterns that aggravate LBP.

Other treatments

The Gpi and STN are used in trials by researchers to observe variations in the length of time that motor improvements last. Parkinsonian symptoms are linked to subthalamic and lenticular nucleus abnormalities. In order to enhance gait control, STN DBS with stimulation of the pedunclopontine nucleus is used. Treatment-resistant PD patients benefit from the ongoing advancement of technologies that alter the brain's circuitry. In order to provide the brain with small amounts of electrical current, electrodes are affixed to the skin or placed beneath it for transcranial direct current stimulation. It could make it possible to postpone beginning DBS treatment and lessen the frequency of drug administration.

ParkinStim is an item worn at night. It can help people with PD wake up feeling better during the night since ParkinStim's symptoms tend to wear off in the morning.

Wireless instantaneous neurotransmitter concentration sensor devices help measure chemical or neurotransmitter release in the brain. This is being utilized to examine brain activity during DBS in combination with fMRI (functional magnetic resonance imaging).

Massage treatments, yoga, tai chi, and physical, occupational, and speech therapies can help relieve stress and improve flexibility. Exercise and a balanced diet aid in developing muscles and enhance balance, flexibility, and coordination.

Glial cells, too, formed from neurotrophic factors, are a source of protein that supports and bolsters DA-producing cells. Enzyme inhibitors (such as COMT and MAD-B inhibitors) boost DA by slowing down the enzymes that break down DA. Anticholinergic medications decrease tremors and muscle stiffness, whereas amantadine decreases involuntary movements.

Given our present understanding of the pathophysiology of PD, effective neuroprotective medicines that target microglial activation's potentially harmful consequences, reduce oxidative

or proteolytic stress, or encourage neuronal growth and repair are indicated. These medicines ought to be successful in stopping, delaying, or even reversing the PD's natural progression and underlying neurodegenerative process. These treatments include pharmacological interventions that increase intracellular protein clearance or inhibit PCD pathways, antioxidants, anti-inflammatory drugs, neural growth factor infusions, brain transplant operations, and possible gene therapies.

Pharmacologists and gene-based treatments are two experimental techniques that aim to implement the therapy at or before hazardous exposure or the manifestation of harmful alleles that cause genetic mutations that disrupt normal cellular functions. Dopaminergic stimulation, linked to motor symptoms, is a side effect of dopaminergic therapy that includes some discomfort. Thus, treating PD pain entails DA agonists, safinamide, antiparkinson drugs, rotigotine's positive effect on fluctuation-related pain, and the rescue of dopaminergic pathways.

The effects of available analgesics are efficient in reducing various forms of pain in people with PD. Common analgesics such as tramadol, oxycodone, and opioids are complementary therapies for analgesics; however, they are not explicitly made for PD pain.

Moreover, mixed serotonin and noradrenaline reuptake inhibitors, such as duloxetine, are advised for PD patients experiencing both pain and sadness. Non-steroidal anti-inflammatory medicines (NSAIDs), such as acetaminophen, naproxen, and ibuprofen for the general public, are other therapies for PD. These can be well tolerated because they do not have any adverse neurological effects, but they may have others. A process termed an epidural injection, which reduces swelling and inflammation to manage pain, may be mentioned by some medical professionals. If a particular structural cause is identified, surgery could also be advised.

Diet

Food habits and diets are significant factors in PD; thus, it is advisable to follow an organic diet and be mindful of what PD patients put into their bodies. Given that pesticides and herbicides have been linked to PD cases and higher rates, it is important to be aware of the 1,500-mile transit distance between the food's point of origin and the pesticide usage of farmers. Consuming raw, fresh veggies raises folic acid levels in the body, which lowers the risk of PD. The usual use of omega-3 fatty acids is to stop cell death and degeneration, improve cardiovascular health, increase immunity, and regulate cholesterol levels. In 2008 Canadian research, mice were given omega-3 supplements for ten months. Before being infected with a chemical that may induce PD, another set of mice was maintained as a control. The omega-3 group showed no drop in DA and no symptoms of PD, whereas the control group experienced a sharp decline in DA levels. Vegetables, fruits, nuts, whole grains, olive oil, seeds, and plant-based cuisine are among the things that should be consumed. Eggs, poultry, dairy, and seafood should be consumed in moderation. Red meat and desserts are to be consumed infrequently. Diets heavy in processed meals, rich in citric juices, high in saturated fats, sugary beverages or foods, alcohol, and difficult-to-chew foods are the ones to stay away from. It is also critical to understand that iron may impair the absorption of PD medicines. Dietitians advise

reducing sugar intake, increasing high-fiber foods, and drinking lots of water to support general health. Regular aerobic exercise also helps to counteract the inflammatory signals that contribute to the development of PD by reducing inflammation in the brain.

Green tea reduces PD severity, enhances brain function, is nutrient-dense, anti-inflammatory, and lowers cancer risk. Green tea can maintain DA levels in sick brain tissues and offer other protective effects on the brain's neuronal network. Physicians advise choosing high-quality green tea since certain low-quality varieties include fluoride, which has been linked to cognitive decline. It has also been shown that consistent coffee use slows the death of neurons that produce DA. Similarly, due to its antioxidative properties, uric acid reduces the likelihood of developing several neurological illnesses. Michael J. Fox Foundation's clinical study showed that inosine safely raised uric acid levels to slow PD progression.

Vitamin D

Mineral absorption requires vitamin D3, obtained from sunshine and grass-sourced animal fat. It can strengthen immunity, prevent osteoporosis, promote bone health, guard against cancer and dementia, and elevate mood and energy levels. One of the many detrimental impacts of aging is that, without vitamin D, humans cannot absorb the necessary quantities of calcium and phosphorus for normal function. Low vitamin D levels were present in 70% of early PD patients. Vitamin D is chemically altered and absorbed with cholesterol and vitamins into circulation. Vitamin D supplements may be helpful since higher amounts of vitamin D are linked to lower incidences of PD.

Ongoing Research

Integrating toxic-induced and genetic research models can provide a more profound knowledge of the interaction between genetic and environmental variables. By stopping and lowering the death of neurons, researchers can create neuroprotective medicines to prevent or decrease the course of neuroinflammatory diseases. In a paradigm where mice harboring SNCA mutations experienced an adult-onset degenerative illness marked by impaired mobility and α -synuclein aggregation but no dopaminergic neuron loss. Some have LRRK2 mutations yet show no signs of PD symptoms. Though not fully reflected, fruit flies and worms—genetically modified to overexpress LRRK2—have decreased motor function and DA neuron loss. In the first toxic-induced model, DA-producing neurons in the SNc were killed using MPTP, or neurotoxic 6-hydroxy-DA.

Toxins are utilized to investigate how toxins affect mitochondrial function. Herbicides like paraquat and maneb and the insecticide rotenone have beneficial uses for researching the effects of DA depletion. Using induced pluripotent stem cells (iPSCs), researchers created a novel modeling technique for genetic engineering for PD. By modulating skin, blood, hair follicles, or other tissues, researchers can create iPSCs, which can be genetically engineered to produce dopaminergic neurons. This facilitates research into PD's cellular and molecular causes

and prospective cures. The prospect of individualized treatment strategies is made possible by the capacity to generate neurons in PD patients.

In PD research, biomarkers (BM) are the holy grail. BM are observable, quantifiable alterations in the body that may be used to forecast, identify, and track the onset and course of the disease. Numerous techniques can be used to identify BM, including genetic research, biological samples (spinal and cerebrospinal fluid), and imaging scans. Perfect PD BMs are simple to test, change in intensity according to the disease, and are aberrant in the preclinical stage of the illness before symptoms appear. Credible BMs enable doctors to identify patients more likely to acquire PD and precisely track the course of the illness.

The FDA has approved the evaluation of people with suspected parkinsonism using brain imaging technologies to identify DA transporters. DaTscan detects if there has been a loss of DA-producing neurons using single-photon emission computed tomography and iodine-based radioactive chemicals. However, it cannot reliably differentiate this condition from other illnesses. Furthermore, the development of positron emission tomography imaging agents that demonstrate the buildup of α -synuclein in the brain is necessary. This might monitor the level of α -synuclein and offer a way to assess whether treatments targeted at lowering α -synuclein are effective or ineffective. To target α -synuclein, a radiotracer agent injection is necessary.

According to NINDS research that focuses on possibly creating neuroprotective medicines that adjust calcium levels for individuals who have just been diagnosed, DA neurons and body cells maintain optimal amounts of calcium flowing through the channels. Cells malfunction when calcium levels are too low and perish when levels are too high, contributing to the onset of PD. This suggests that using calcium channel blockers to modulate the effects of calcium slows the course of PD. Potentially, among animal models, isradipine—a calcium channel blocker—shields DA neurons against a range of poisons.

Conclusion

In conclusion, Parkinson's disease poses a significant global health challenge, necessitating a comprehensive understanding of its etiology, pathophysiology, diagnosis, and treatment. As the neurotransmitter dopamine plays a crucial role in the development and progression of PD, the decrease, death, and loss of dopamine in the substantia nigra have been the primary cause of PD. Defining motor deficits in PD is linked to the vulnerability of specific subgroups in nigral DA neurons. As the neurons in the brain weaken or get damaged, patients with PD will experience difficulty in movement. Although the cause of PD remains uncertain and is not currently known to have treatments that delay the disease's course or have a neuroprotective impact, treating the motor and nonmotor symptoms of PD is the main aim of the disease management strategy, which also aims to enhance the patient's overall quality of life.

Works Cited

- Butch, R., Wasta, V., & Tucker, A. (2022, May 19). ‘Sting’ protein’s efforts to clean up brain cell damage may speed Parkinson’s disease progress. *Johns Hopkins Medicine*.
<https://www.hopkinsmedicine.org/news/newsroom/news-releases/2022/05/sting-proteins-efforts-to-clean-up-brain-cell-damage-may-speed-parkinsons-disease-progress>
- Lanciego, J. L., Luquin, N., & Obeso, J. A. (2012). Functional neuroanatomy of the basal ganglia. *Cold Spring Harbor Perspectives in Medicine*, 2(12), a009621.
<https://doi.org/10.1101/cshperspect.a009621>
- Blandini, F., Nappi, G., Tassorelli, C., & Martignoni, E. (2000). Functional changes of the basal ganglia circuitry in Parkinson’s disease. *Progress in Neurobiology*, 62(1), 63–88.
[https://doi.org/10.1016/s0301-0082\(99\)00067-2](https://doi.org/10.1016/s0301-0082(99)00067-2)
- Bsn, R. Z. R. (2023, February 8). *What part of your brain does Parkinson’s disease affect?* Healthline.
<https://www.healthline.com/health/parkinsons-disease/parkinsons-disease-affects-brain>
- Zhou, C., Guan, X., Guo, T., Zeng, Q., Gao, T., Huang, P., Xuan, M., Gu, Q., Xu, X., & Zhang, M. (2019). Progressive brain atrophy in Parkinson’s disease patients who convert to mild cognitive impairment. *CNS Neuroscience & Therapeutics*, 26(1), 117–125.
<https://doi.org/10.1111/cns.13188>
- Robbins, T. W., & Everitt, B. J. (1992). Functions of dopamine in the dorsal and ventral striatum. *Seminars in Neuroscience*, 4(2), 119–127. [https://doi.org/10.1016/1044-5765\(92\)90010-y](https://doi.org/10.1016/1044-5765(92)90010-y)
- Braak, H., Braak, E., Yilmazer, D., De Vos, R. a. I., Jansen, E. N. H., Bohl, J., & Jellinger, K. A. (1994). Amygdala pathology in Parkinson’s disease. *Acta Neuropathologica*, 88(6), 493–500. <https://doi.org/10.1007/bf00296485>
- Smith, Y., Villalba, R. M., & Raju, D. V. (2009). Striatal spine plasticity in Parkinson’s disease: pathological or not? *Parkinsonism & Related Disorders*, 15, S156–S161.
[https://doi.org/10.1016/s1353-8020\(09\)70805-3](https://doi.org/10.1016/s1353-8020(09)70805-3)
- Namkung, H., Kim, S. H., & Sawa, A. (2017). The Insula: an underestimated brain area in clinical neuroscience, psychiatry, and neurology. *Trends in Neurosciences*, 40(4), 200–207. <https://doi.org/10.1016/j.tins.2017.02.002>
- Fathy, Y. Y., Hepp, D. H., De Jong, F. J., Geurts, J. J. G., Foncke, E., Berendse, H. W., Van De Berg, W. D. J., & Schoonheim, M. M. (2020). Anterior insular network disconnection and cognitive impairment in Parkinson’s disease. *NeuroImage: Clinical*, 28, 102364.
<https://doi.org/10.1016/j.nicl.2020.102364>
- Mokhtar, M., & Singh, P. (2023, July 24). *Neuroanatomy, periaqueductal gray*. StatPearls - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK554391/>
- Buhidma, Y., Hobbs, C., Malcangio, M., & Duty, S. (2023). Periaqueductal grey and spinal cord pathology contribute to pain in Parkinson’s disease. *Npj Parkinson’s Disease*, 9(1).
<https://doi.org/10.1038/s41531-023-00510-3>
- Hindle, J. V. (2010). Ageing, neurodegeneration and Parkinson’s disease. *Age And Ageing*, 39(2), 156–161. <https://doi.org/10.1093/ageing/afp223>

- The genetic link to Parkinson's disease.* (2022, April 10). Johns Hopkins Medicine.
<https://www.hopkinsmedicine.org/health/conditions-and-diseases/parkinsons-disease/the-genetic-link-to-parkinsons-disease>
- Pedersen, T. (2023, April 24). *Is alcohol a risk factor for Parkinson's disease?* Healthline.
<https://www.healthline.com/health/parkinsons/parkinsons-and-alcohol>
- Parkinson's Disease: causes, symptoms, and treatments.* (n.d.). National Institute on Aging.
<https://www.nia.nih.gov/health/parkinsons-disease/parkinsons-disease-causes-symptoms-and-treatments#symptoms>
- Gilbert, R., MD PhD. (2021, August 24). *Lower back pain and Parkinson's disease.* American Parkinson Disease Association.
<https://www.apdaparkinson.org/article/lower-back-pain-and-parkinsons-diseaselower-back-pain-and-parkinsons-disease/>
- Living, J. S. (2022, October 28). *7 Scientifically backed ways to prevent Parkinson's Disease.* Judson.
<https://www.judsonsmartliving.org/blog/7-scientifically-backed-ways-to-prevent-parkinsons-disease/>
- Padmini, P., Hickey, F., Patel, L., & Sannar, E. M. (2018). Levodopa-Responsive Early-Onset parkinsonism in Down Syndrome. *Case Reports in Neurological Medicine*, 2018, 1–3.
<https://doi.org/10.1155/2018/2314791>
- Parkinson's Disease: challenges, progress, and promise.* (n.d.). National Institute of Neurological Disorders and Stroke.
<https://www.ninds.nih.gov/current-research/focus-disorders/parkinsons-disease-research/parkinsons-disease-challenges-progress-and-promise>
- Alexander, G. E. (2004). Biology of Parkinson's disease: pathogenesis and pathophysiology of a multisystem neurodegenerative disorder. *Dialogues in Clinical Neuroscience*, 6(3), 259–280. <https://doi.org/10.31887/dcns.2004.6.3/galexander>

The Journey to Alzheimer's Disease Diagnosis and Beyond Through Different Perspectives

By Deetya Potakamuri

Abstract

Alzheimer's disease (AD) is a prevalent neurodegenerative disease that affects at least 50 million people worldwide and is on the rise to affect about 152 million by 2050. This progressive disease burdens both the patients and their family/caregiver, bringing upon social, emotional, and financial challenges. The pathology of AD involves the build up of amyloid plaques and neurofibrillary tangles years before the onset and diagnosis of the disease. Signs of cognitive decline are often first reported to health care professionals before evaluation and eventual diagnosis and treatment. Initially, these signs are often diagnosed as mild cognitive impairment (MCI), but consistent further testing is required by health care professionals to monitor the progress of the disorder and its possible development into dementia and AD. Health care professionals utilize a combination of cognitive assessment, evidence of AD pathology, and other testing to diagnose the disease and rule out other potential causes of cognitive impairment. An earlier diagnosis of AD is beneficial in planning current and future care management for AD symptoms as well as navigating the limitations for the patient, their families/caregivers, and the health care professionals themselves. In this review article, the current etiology of AD will be discussed in addition to the methods of diagnosis and treatments for the disease. Additionally, the perspectives of health care providers, patients and their family, and society as a whole will be considered in the future of AD.

Introduction

Dementia, including Alzheimer's Disease (AD), which is a progressive neurodegenerative disease most common in adults over the age of 65, currently affects at least 55 million people worldwide, and their incidence is projected to rise to 152 million affected by 2050 (GBD 2019 Dementia Forecasting Collaborators, 2022; World Health Organization, 2023). This projected increase can be attributed to an aging population, aided by the rising life expectancy, forecasted to grow by over four years in both men and women by 2050 (Stein Emil Vollset et al., 2024). AD is the most common type of dementia, with 60-70% of all dementia cases (World Health Organization, 2023). Currently, there is no cure for AD, and the available treatments focus on managing the symptoms of the disease. Some newer treatments, like the drugs lecanemab and aducanumab, have been shown to slow cognitive decline from AD by targeting the disease's underlying pathology (Alzheimer's Association, 2024; FDA, 2023; Sevigny et al., 2016). Such novel developments in AD research are promising, and such interventions can help reduce the prevalence and burden of the disease (Brookmeyer et al., 1998). However, for any such intervention to be effective, it must fit within the emotional, physical, and financial interests of AD patients and their support systems, including health care providers and caregivers.

In this review article, the etiology of AD will be covered, specifically examining the disease's progression from the preclinical stage to the later stages of worsening dementia until death. The current methods and criteria for diagnosing AD and the latest treatment options to manage the disease will be reviewed. Finally, the real-life impact of AD/dementia through the perspectives of health care professionals, affected families, and American society at large will be investigated, which should be considered when developing and implementing new interventions and guidelines for combating AD as the century progresses.

Progression of Alzheimer's Disease

Preclinical Alzheimer's Disease

In the preclinical stage of AD, amyloid beta protein, whose presence is characteristic of AD, can start building up as early as 20-30 years before the onset of dementia (Jansen et al., 2015). Additionally, the levels of phosphorylated tau protein, another hallmark of AD, rise to high levels as amyloid plaques build up in the brain (Barthélemy et al., 2020; Wegmann et al., 2021). Although amyloid buildup to a degree isn't unusual in cognitively normal older adults, amyloid levels can grow to an unhealthy amount in AD patients (Jansen et al., 2015). Cognitive decline is associated with this increase in amyloid plaques, especially in the more advanced stages within preclinical AD with unusual amyloid levels, injury markers, and slight changes in one's cognitive abilities (Vos et al., 2013). Elevated amyloid levels in older individuals have been associated with worse cognitive results on the Preclinical Alzheimer's Cognitive Composite test compared to those without higher amyloid levels (Sperling et al., 2020).

In a study estimating the duration of the different stages of AD considering age, gender, and the APOE4 ϵ 4 carrier status, the preclinical stage of AD in most of the scenarios is estimated to be the longest, especially in a research setting where AD and its progression is studied (Vermunt, 2019). This cohort study of people with amyloid pathology, where participants were studied through their lives, estimates that 20% of those with preclinical AD progress to mild cognitive impairment (MCI) due to AD after a median 3.8 year follow-up (Vermunt, 2019). Because only about a fifth of people with preclinical AD go on to develop Alzheimer's Disease, it can be difficult to diagnose AD with certainty in the early stages. The fact that older adults, who are at the most risk of developing AD, are more likely to have amyloid buildup even without the disease further complicates early diagnosis.

Mild Cognitive Impairment

MCI, a stage of cognitive decline not normal for one's age but not classified as dementia, is clinically diagnosed due to an observed change in cognition, impairment in at least one cognitive domain, and struggle to perform certain complex tasks efficiently yet able to maintain their independence (Albert et al., 2011). The cognitive domains that can be impacted include language, executive function, visuospatial processing, and memory, with impairments in episodic memory being commonly observed in those who progress further to dementia due to AD (*A Healthcare Provider's Guide to Mild Cognitive Impairment (MCI): Diagnosis, Pharmacologic*

Management, Non-Pharmacologic Management, and Other Considerations, n.d.; Chehretnegar et al., 2021; McCullough et al., 2019). It is estimated that 45% of MCI patients progress to AD mild dementia, according to a median 3.9-year follow-up (Vermunt, 2019). Therefore, at the MCI stage, it is twice as likely that the patient progresses to the next stage in the next few years than patients with preclinical AD to MCI. The MCI stage of AD may be a more optimal stage for AD intervention due to the greater certainty of AD in addition to the observation that those with MCI are still able to retain their independence despite some cognitive challenges.

Mild Dementia

In this early stage of AD dementia, the patient has progressed to dementia. Although affected individuals may face difficulties in higher-level tasks like personal financial management and act slower in general while completing other tasks, they are still able to maintain their independence and manage most of their lives on their own (Alzheimer's Association, 2024). Individuals may need some support at this stage from trusted close people to keep up with some responsibilities, including financial as the person affected is likely dealing with retirement and now planning for their future with AD. An estimated 30% of mild dementia patients progress further to moderate and severe dementia due to AD after a median two-year follow-up (Vermunt, 2019). The follow up period being shorter compared to the preclinical and MCI stages can be attributed to the shorter durations of later AD stages compared to early stages of the disease (Vermunt, 2019). Since the individual has dementia at this point that can be disruptive to the flow of their lives, the current patient quality of life should become more of a focus when planning disease treatment and management.

Moderate/Severe Dementia

As AD progresses and further damage accumulates in the brain, impacted individuals greatly lose their independence from not being able to do daily self-care tasks such as going to the bathroom or bathing, placing the burden on their caregivers to provide around-the-clock care as their brain and body functions progressively shut down in their final stretch of life (Alzheimer's Association, 2024). The whole body is now impacted and vulnerable as it prepares for death. Important body processes like swallowing become more difficult and hazardous, a dysfunction known as dysphagia for late-stage AD patients in particular (Homer et al., 1994; Johns Hopkins Medicine, 2024). Dysphagia contributes to the development of the most frequent cause of death in AD patients, aspiration pneumonia (Kalia, 2003). Overall, Alzheimer's disease is a major cause of death and was the 7th leading cause of death in the United States in years following the start of the COVID pandemic (Centers for Disease Control and Prevention, 2021).

Diagnosis

Clinical Diagnosis

Through a combination of different tests and evaluations conducted over a period of time, a diagnosis of AD can be made. Health care providers may conduct various cognitive tests, order

blood tests, analyze cerebral spinal fluid (CSF), perform psychiatric evaluations, and conduct supportive brain scans on patients with reported cognitive symptoms (National Institute on Aging, 2022). In an updated version of the 1984 NINCDS-ADRDA criteria for AD dementia diagnosis by a workgroup established by the National Institute on Aging and the Alzheimer's Association, the characteristics of a probable or possible AD dementia diagnosis are outlined for consideration in a clinical setting (McKhann et al., 2011). For a patient to be diagnosed with probable AD dementia, the criteria include the gradual onset of symptoms, short term memory loss, cognitive impairment in one other cognitive domain besides learning and memory, and the absence of events like strokes from cerebrovascular disease that can also be responsible for cognitive impairment (McKhann et al., 2011). The differential diagnosis of possible AD dementia includes the characteristics of either a more sudden change in cognitive impairment or the presence of other contributing factors, such as a history of stroke or other neurological diseases (McKhann et al., 2011). Due to the many factors that must be considered when diagnosing AD with a certain degree of certainty, the diagnosis journey can be lengthy and costly with the amount of tools and testing used in evaluation for diagnosis. Some of the criteria for diagnosis like the gradual onset of symptoms needs to be observed by patients, their family, and health providers and noted as time passes to demonstrate the criteria. If the patient isn't able to notice these changes and doesn't have family that interact with them daily, meeting such criteria can become more difficult and the initial visit to the doctor with concern to the developing symptoms may be delayed, further slowing down the diagnosis of AD.

Treatments

At present, there are several FDA-approved drugs that either target the disease progression of AD or manage its cognitive and behavioral symptoms (*FDA-Approved Treatments for Alzheimer's*, n.d.). Additionally, non-drug interventions, such as orientation training, which help improve a patient's sense of space and time, and various mental exercises, may also be utilized if deemed suitable (National Library of Medicine, 2022).

Disease Modifying Therapies

Currently, many therapies target AD pathology, classified as disease-modifying therapies, in phases 2 and 3 of clinical testing (Cummings et al., 2019). In July 2024, the U.S. Food and Drug Administration approved Lilly's Kisunla™ (donanemab-azbt) injection as a treatment for patients with mild cognitive impairment (MCI) or mild dementia (Research, 2024). In a randomized clinical trial conducted by Eli Lilly and Company on those with early symptomatic AD, the participants with early symptomatic AD who received the donanemab had lower levels of cognitive impairment than those who received placebo (Sims et al., 2023). The side effects of Kisunla include amyloid-related imaging abnormalities (ARIA), which can lead to temporary brain swelling that resolves on its own without major complications in most cases and was seen to be 10x more common in those who received donanemab in the Lilly clinical trial (Research, 2024; Sims et al., 2023). Lecanemab, sold as Leqembi™ by Eisai, was approved by the FDA a

year earlier, is another treatment that targets the underlying AD pathology and was confirmed in a randomized phase 3 clinical trial to benefit recipients of the treatment by reducing amyloid plaque build up in the brain, particularly in those with MCI or mild dementia (*CTG Labs - NCBI*, n.d.; FDA, 2023; “*LEQEMBI®*” (*Lecanemab*) *Approved for the Treatment of Alzheimer’s Disease in China* | *News Release: 2024*, n.d.; Van Dyck et al., 2022). A third drug that targets AD pathology, aducanumab, is set to be discontinued in 2024 for management reasons by the company Biogen, the drug manufacturer, without any concerns to safety (Alzheimer’s Association, 2024; Research, 2023; U.S. Food and Drug Administration, 2021). This factor of the drug industry and the role of business is important to consider due to the capitalized health care system. In the case of aducanumab, the drug was discontinued for business-related reasons rather than health-related reasons, demonstrating the significance of money in AD disease care not only for those directly and personally affected, but also for the healthcare industry whose aim is to sell health care at a favorable profit. Focusing on the new treatments, in order to maximize the effectiveness of these newer disease modifying treatments, an earlier diagnosis of AD is necessary to target and manage the disease before its symptoms and damage permanently worsen as AD progresses to the later stages.

Treating Symptoms

The drugs approved by the FDA for the management of the symptoms of Alzheimer’s disease include donepezil, galantamine, memantine, and rivastigmine (*FDA-Approved Treatments for Alzheimer’s*, n.d.; Howard et al., 2012; Raskind et al., 2004). Three of the drugs, donepezil, galantamine, and rivastigmine, are the cholinesterase inhibitors that stop acetylcholine, an important messenger involved in nerve cell communication, from breaking down (*FDA-Approved Treatments for Alzheimer’s*, n.d.). The exposure to cholinesterase inhibitors have been shown to reduce the mortality among dementia-prone groups, particularly nursing home residents (Havreng-Théry et al., 2024).

Additionally, the drugs brexpiprazole and suvorexant are FDA-approved to manage the behavioral and psychological symptoms of AD (*FDA-Approved Treatments for Alzheimer’s*, n.d.; Grossberg et al., 2024; Herring et al., 2020). Brexpiprazole has been approved as the Rexulti™ supplement and has been shown to manage the higher prevalence of agitation in AD patients, which includes restlessness, hostile language, and physical aggression (Commissioner, 2023; Lee et al., 2023; Van der Mussele et al., 2014). The other drug used to treat the non-cognitive symptoms of AD is suvorexant, which is administered through Belsomra™’s tablets and is utilized to manage the sleep disturbance of insomnia in patients whose AD disease progression is in the mild to moderate stages (Herring et al., 2020; *HIGHLIGHTS of PRESCRIBING INFORMATION*, n.d.; *Merck Receives Approval for BELSOMRA® (Suvorexant) C-IV Label Update to Include Findings from Study of Insomnia in Patients with Mild-To-Moderate Alzheimer’s Disease*, n.d.). Suvorexant is an orexin receptor antagonist that blocks OX1R and OX2R, which both contribute to wakefulness (Bennett et al., 2014).

Perspectives

Health Care Provider

Health care providers, who are responsible for maintaining the health of their patients through diagnosis, prevention, and treatment, are an important party to consider in improving the outcomes of people with AD, especially in the early diagnosis and treatment of AD (Monica Zigman Suchsland et al., 2023; World Health Organization, 2013). In interviews with ten health care providers to gather their insight and opinion on the KAER (Kickstart, Assess, Evaluate, Refer) toolkit, which aims to improve the process of diagnosing AD, although they found the toolkit to be useful, they also emphasized how such methods of cognitive assessment must be easy to use and efficient within their means (Monica Zigman Suchsland et al., 2023). Although health care providers have a critical role in the outcomes of their patients, they are often restricted by the health care systems they work in and may not always be able to provide the necessary level or type of care (Sideman et al., 2023). For instance, 69% of health care professionals acknowledge that in medical school, the condition of dementia was taught to a very limited degree or not at all despite older patients' interest on that topic as they are getting older (Alzheimer's Impact Movement, 2024). Coupled with the general lack of teaching of dementia in standard medical education, is the shortage of dementia specialists, impacting U.S. states like Kansas (Alzheimer's Association, 2024; Alzheimer's Impact Movement, 2024).

Patient and Family

The journey through AD and the major life changes associated with its occurrence affects both patients and their families significantly, particularly women, who consist of two-thirds of caregivers for people with AD (Alzheimer's Association, 2024). Additionally, according to a case-control study in Northern Italy, women were shown to have a higher risk of developing AD compared to men, although the risk of developing dementia was similar among both genders (Adani et al., 2020). On average, caregivers provide 31 hours of care to dementia patients per week, dedicating about a fifth of their lives to caregiving (Alzheimer's Association, 2024). Such long hours and the trauma of witnessing the progressive mental decline of loved ones have been shown to drastically affect the mental health of dementia caregivers, with 32.3% of caregivers experiencing suicide ideation compared to the 4.3% of adults annually on average in the United States (Ivey-Stephenson et al., 2022; Solimando et al., 2022). Therefore, it is important to consider family, the main support system of a patient, when designing and implementing new treatments and caregiving practices. Prioritizing the health of the loved ones who stand by and support their sick family member is necessary to maximize human health and emotional well-being as well as promote the bonds of relationships in times of sickness.

Society

According to the Alzheimer's Association's 2024 *Alzheimer's Disease Facts and Figures* report, a projected \$360 billion will be spent on the healthcare and long-term care of those with Alzheimer's or other dementias in the United States in 2024 (Alzheimer's Association, 2024).

\$231 billion of that sum will be spent by the government programs of Medicare and Medicaid (Alzheimer's Impact Movement, 2024). If there are no major developments in dementia care or treatment, costs of this disease will increase to almost \$1 trillion in 2050, increasing both out-of-pocket and government spending under Medicare and Medicaid (Alzheimer's Association, 2024). These staggering costs could potentially bankrupt Medicare and Medicaid, especially with the projected increase in AD and other dementias in the next few decades (Alzheimer's Association, 2014; GBD 2019 Dementia Forecasting Collaborators, 2022). This increase in dementia can be attributed to the aging world population, which is caused by a greater life spans due to medical advancements and in the coming decades, will lead to a greater proportion of older, retired individuals compared to young, working people (GBD 2019 Dementia Forecasting Collaborators, 2022; United Nations, 2019). A demographic shift to older population from younger population could mean decreased government tax revenues that are utilized to fund public health care services like Medicare and Medicaid. This is particularly problematic when the need for such services will likely go up with the coming rise in AD prevalence.

However, focusing on the early diagnosis of AD has been predicted to lead to a better prognosis of the disease from an economic standpoint for health care costs. In the scenario where AD is diagnosed with greater frequency in the mild cognitive impairment (MCI) stage, it is projected that there will be \$7 trillion in cumulative savings in the United States which \$5.6 trillion would be saved in Medicare and Medicaid costs from the early diagnosis among those alive in 2018 who go on to develop AD (Alzheimer's Association, 2018). This includes medical, pharmaceutical, and long-term care costs (Alzheimer's Association, 2018; Medicare.gov, 2019; *Prescription Drug Coverage*, 2019). Therefore, working towards earlier diagnoses of AD is essential for lowering health care costs for Americans with AD and reducing the cost burden on crucial American health insurance programs. Early AD diagnosis can promote a more controlled and sustainable health care system that is better able to manage and take care of the health needs of people with AD and others in need of care in the future.

Conclusion

In this review article, the perspectives and interests ranging from that of healthcare providers to American society as a whole on Alzheimer's Disease were covered. Additionally, the fact that the current number of health care providers, who already face burnout, who diagnose and treat AD patients aren't sufficient was highlighted, especially with the doubling and tripling of dementia cases in the US and worldwide in the coming decades (Centers for Disease Control and Prevention, 2023). The importance of diagnosing and treating AD earlier on in prognosis was also emphasized because the ability of certain novel treatments to slow down disease progression is particularly more effective in the earlier stages of prognosis. Therefore, despite current dementia medical provider limitations, it is essential that steps are taken for the timely diagnosis and treatment of AD, especially as new treatments gain traction and are utilized.

Works Cited

- Adani, G., Filippini, T., Garuti, C., Malavolti, M., Vinceti, G., Zamboni, G., Tondelli, M., Galli, C., Costa, M., Vinceti, M., & Chiari, A. (2020). Environmental Risk Factors for Early-Onset Alzheimer's Dementia and Frontotemporal Dementia: A Case-Control Study in Northern Italy. *International Journal of Environmental Research and Public Health*, 17(21), 7941. <https://doi.org/10.3390/ijerph17217941>
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270–279. <https://doi.org/10.1016/j.jalz.2011.03.008>
- Alzheimer's Association. (2024). Aducanumab Approved for Treatment of Alzheimer's Disease. *Alzheimer's Disease and Dementia*. <https://www.alz.org/alzheimers-dementia/treatments/aducanumab>
- Alzheimer's Association. (2018). 2018 Alzheimer's Disease Facts and Figures. *Alzheimer's Association*. <https://www.alz.org/media/HomeOffice/Facts%20and%20Figures/facts-and-figures.pdf>
- Alzheimer's Association. (2024). 2024 Alzheimer's Disease Facts and Figures. *Alzheimer's Association*. <https://www.alz.org/media/Documents/alzheimers-facts-and-figures.pdf>
- Alzheimer's Association. (2014). The most expensive disease: Alzheimer's disease. https://act.alz.org/site/DocServer/112114_Most_Expensive_Disease_Handout.pdf?docID=40161
- Alzheimer's Impact Movement. (2024). Costs of Alzheimer's to Medicare and Medicaid. *Alzheimer's Impact*. <https://portal.alzimpact.org/media/serve/id/62509c7a54845>
- Alzheimer's Impact Movement. (2024). Dementia-Capable Workforce. *Alzheimer's Impact*. <https://portal.alzimpact.org/media/serve/id/62509c7a54845>
- Barthélemy, N. R., Li, Y., Joseph-Mathurin, N., Gordon, B. A., Hassenstab, J., Benzinger, Tammie, L. S., Buckles, V., Fagan, A. M., Perrin, R. J., Goate, A. M., Morris, J. C., Karch, C. M., Xiong, C., Allegri, R., Mendez, P. C., Berman, S. B., Ikeuchi, T., Mori, H., Shimada, H., & Shoji, M. (2020). A soluble phosphorylated tau signature links tau, amyloid and the evolution of stages of dominantly inherited Alzheimer's disease. *Nature Medicine*, 26(3), 398–407. <https://doi.org/10.1038/s41591-020-0781-z>
- Bennett, T., Bray, D., & Neville, M. W. (2014). Suvorexant, a Dual Orexin Receptor Antagonist for the Management of Insomnia. *Pharmacy and Therapeutics*, 39(4), 264–266. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3989084/>
- Brookmeyer, R., Gray, S., & Kawas, C. (1998). Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *American Journal of Public Health*, 88(9), 1337–1342. <https://doi.org/10.2105/ajph.88.9.1337>

- Centers for Disease Control and Prevention. (2021). Demographic characteristics of people vaccinated against COVID-19: United States, December 2020–September 2021 (NCHS Data Brief No. 492). <https://www.cdc.gov/nchs/data/databriefs/db492-tables.pdf#4>
- Centers for Disease Control and Prevention. (2023, October 24). Health Workers Face a Mental Health Crisis. Centers for Disease Control and Prevention. <https://www.cdc.gov/vitalsigns/health-worker-mental-health/index.html>
- Chehrehnegar, N., Shati, M., Esmaili, M., & Foroughan, M. (2021). Executive function deficits in mild cognitive impairment: evidence from saccade tasks. *Aging & Mental Health*, 1–9. <https://doi.org/10.1080/13607863.2021.1913471>
- Commissioner, O. of the. (2023, May 11). FDA Approves First Drug to Treat Agitation Symptoms Associated with Dementia due to Alzheimer’s Disease. FDA. <https://www.fda.gov/news-events/press-announcements/fda-approves-first-drug-treat-agitation-symptoms-associated-dementia-due-alzheimers-disease>
- CTG Labs - NCBI. (n.d.). Clinicaltrials.gov. <https://clinicaltrials.gov/study/NCT03887455>
- Cummings, J., Lee, G., Ritter, A., Sabbagh, M., & Zhong, K. (2019). Alzheimer’s Disease Drug Development pipeline: 2019. *Alzheimer’s & Dementia: Translational Research & Clinical Interventions*, 5(1), 272–293. <https://doi.org/10.1016/j.trci.2019.05.008>
- FDA. (2023, July 6). FDA Converts Novel Alzheimer’s Disease Treatment to Traditional Approval. FDA. <https://www.fda.gov/news-events/press-announcements/fda-converts-novel-alzheimers-disease-treatment-traditional-approval>
- FDA-approved treatments for Alzheimer’s. (n.d.). <https://www.alz.org/getmedia/4f4ca289-a2c6-4df9-8cdf-390365bd477e/alzheimers-dementia-fda-approved-treatments-for-alzheimers-ts.pdf>
- GBD 2019 Dementia Forecasting Collaborators. (2022). Estimation of the Global Prevalence of Dementia in 2019 and Forecasted Prevalence in 2050: an Analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*, 7(2). <https://pmc.ncbi.nlm.nih.gov/articles/PMC8810394/>
- Grossberg, G., Lee, D., Slomkowski, M., Nanco Hefting, Chen, D., Hobart, M., & Cummings, J. (2024). Safety, Tolerability, and Efficacy of Brexpiprazole for the Treatment of Agitation Associated with Dementia Due to Alzheimer’s Disease: A 12-week Extension Trial (P11-9.011). *Neurology*, 102(17_supplement_1). <https://doi.org/10.1212/wnl.00000000000204882>
- Havreng-Théry, C., Oquendo, B., Zolnowski-Kolp, V., Krolak-Salmon, P., François Bertin-Hugault, Lafuente-Lafuente, C., & Joël Belmin. (2024). Cholinesterase inhibitors and memantine are associated with a reduced mortality in nursing home residents with dementia: a longitudinal observational study. *Alzheimer's Research & Therapy*, 16(1). <https://doi.org/10.1186/s13195-024-01481-0>
- A Healthcare Provider’s Guide to Mild Cognitive Impairment (MCI): Diagnosis, pharmacologic management, non-pharmacologic management, and other considerations. (n.d.).

- https://memory.ucsf.edu/sites/memory.ucsf.edu/files/wysiwyg/UCSF_MCI_Providers_7-13-17.pdf
- Herring, W. J., Ceesay, P., Snyder, E., Bliwise, D., Budd, K., Hutzelmann, J., Stevens, J., Lines, C., & Michelson, D. (2020). Polysomnographic assessment of suvorexant in patients with probable Alzheimer’s disease dementia and insomnia: a randomized trial. *Alzheimer’s & Dementia*, 16(3), 541–551. <https://doi.org/10.1002/alz.12035>
- HIGHLIGHTS OF PRESCRIBING INFORMATION. (n.d.).
https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/204569s006lbl.pdf
- Homer, J., Alberts, M. J., Dawson, D. V., & Cook, G. M. (1994). Swallowing in Alzheimer’s Disease. *Alzheimer Disease & Associated Disorders*, 8(3), 177–189.
<https://doi.org/10.1097/00002093-199408030-00004>
- Howard, R., McShane, R., Lindesay, J., Ritchie, C., Baldwin, A., Barber, R., Burns, A., Denning, T., Findlay, D., Holmes, C., Hughes, A., Jacoby, R., Jones, R., Jones, R., McKeith, I., Macharouthu, A., O’Brien, J., Passmore, P., Sheehan, B., & Juszczak, E. (2012). Donepezil and Memantine for Moderate-to-Severe Alzheimer’s Disease. *New England Journal of Medicine*, 366(10), 893–903. <https://doi.org/10.1056/nejmoa1106668>
- Ivey-Stephenson, A. Z., Crosby, A. E., Hoenig, J. M., Gyawali, S., Park-Lee, E., & Hedden, S. L. (2022). Suicidal Thoughts and Behaviors Among Adults Aged ≥18 Years — United States, 2015–2019. *MMWR. Surveillance Summaries*, 71(1), 1–19.
<https://doi.org/10.15585/mmwr.ss7101a1>
- Jansen, W. J., Ossenkoppele, R., Knol, D. L., Tijms, B. M., Scheltens, P., Verhey, F. R. J., Visser, P. J., Aalten, P., Aarsland, D., Alcolea, D., Alexander, M., Almdahl, I. S., Arnold, S. E., Baldeiras, I., Barthel, H., van Berckel, B. N. M., Bibeau, K., Blennow, K., Brooks, D. J., & van Buchem, M. A. (2015). Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia. *JAMA*, 313(19), 1924. <https://doi.org/10.1001/jama.2015.4668>
- Johns Hopkins Medicine. (2024). Swallowing disorders. www.hopkinsmedicine.org.
<https://www.hopkinsmedicine.org/health/conditions-and-diseases/swallowing-disorders>
- Kalia, M. (2003). Dysphagia and aspiration pneumonia in patients with Alzheimer’s disease. *Metabolism*, 52, 36–38. [https://doi.org/10.1016/s0026-0495\(03\)00300-7](https://doi.org/10.1016/s0026-0495(03)00300-7)
- Lee, D., Slomkowski, M., Hefting, N., Chen, D., Larsen, K. G., Kohegyi, E., Hobart, M., Cummings, J. L., & Grossberg, G. T. (2023). Brexpiprazole for the Treatment of Agitation in Alzheimer Dementia: A Randomized Clinical Trial. *JAMA Neurology*, 80(12). <https://doi.org/10.1001/jamaneurol.2023.3810>
- “LEQEMBI®” (Lecanemab) Approved for the Treatment of Alzheimer’s Disease in China | News Release: 2024. (n.d.). Eisai Co., Ltd.
<https://www.eisai.com/news/2024/news202403.html>
- McCullough, K. C., Bayles, K. A., & Bouldin, E. D. (2019). Language Performance of Individuals at Risk for Mild Cognitive Impairment. *Journal of Speech, Language, and Hearing Research*, 62(3), 706–722. https://doi.org/10.1044/2018_jslhr-l-18-0232

- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263–269.
<https://doi.org/10.1016/j.jalz.2011.03.005>
- Medicare.gov. (2019). Hospice Care Coverage. Medicare.gov.
<https://www.medicare.gov/coverage/hospice-care>
- Merck Receives Approval for BELSOMRA® (suvorexant) C-IV Label Update to Include Findings from Study of Insomnia in Patients with Mild-to-Moderate Alzheimer's Disease. (n.d.). Merck.com.
<https://www.merck.com/news/merck-receives-approval-for-belsomra-suvorexant-c-iv-label-update-to-include-findings-from-study-of-insomnia-in-patients-with-mild-to-moderate-alzheimers-disease/>
- Monica Zigman Suchsland, Gaster, B., Raetz, J., Belza, B., McGuire, L. C., Olivari, B., Tracy, K., & Fitzpatrick, A. L. (2023). Developing a cognitive assessment toolkit for primary care: qualitative assessment of providers' needs and perceptions of usability in clinical practice. *BMC Health Services Research*, 23(1).
<https://doi.org/10.1186/s12913-023-09991-7>
- National Institute on Aging. (2022, December 8). How is Alzheimer's disease diagnosed? National Institute on Aging.
<https://www.nia.nih.gov/health/alzheimers-symptoms-and-diagnosis/how-alzheimers-disease-diagnosed>
- National Library of Medicine. (2022, March 29). Non-drug interventions for Alzheimer's disease. Nih.gov; Institute for Quality and Efficiency in Health Care (IQWiG).
<https://www.ncbi.nlm.nih.gov/books/NBK279355/>
- Prescription Drug Coverage. (2019). Medicare.gov.
<https://www.medicare.gov/coverage/prescription-drugs-outpatient>
- Raskind, M. A., Peskind, E. R., Truyen, L., Kershaw, P., & Damaraju, C. V. (2004). The Cognitive Benefits of Galantamine Are Sustained for at Least 36 Months. *Archives of Neurology*, 61(2), 252. <https://doi.org/10.1001/archneur.61.2.252>
- Research, C. for D. E. and. (2023). Drug Trials Snapshots: ADUHELM. FDA.
<https://www.fda.gov/drugs/drug-approvals-and-databases/drug-trials-snapshots-aduhelm>
- Research, C. for D. E. and. (2024). FDA approves treatment for adults with Alzheimer's disease. FDA.
<https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-treatment-adults-alzheimers-disease>
- Sevigny, J., Chiao, P., Bussière, T., Weinreb, P. H., Williams, L., Maier, M., Dunstan, R., Salloway, S., Chen, T., Ling, Y., O'Gorman, J., Qian, F., Arastu, M., Li, M., Chollate, S.,

- Brennan, M. S., Quintero-Monzon, O., Scannevin, R. H., Arnold, H. M., & Engber, T. (2016). The antibody aducanumab reduces A β plaques in Alzheimer's disease. *Nature*, 537(7618), 50–56. <https://doi.org/10.1038/nature19323>
- Sideman, A. B., Ma, M., Hernandez de Jesus, A., Alagappan, C., Razon, N., Dohan, D., Chodos, A., Al-Rousan, T., Alving, L. I., Segal-Gidan, F., Rosen, H., Rankin, K. P., Possin, K. L., & Borson, S. (2023). Primary Care Practitioner Perspectives on the Role of Primary Care in Dementia Diagnosis and Care. *JAMA Network Open*, 6(9), e2336030. <https://doi.org/10.1001/jamanetworkopen.2023.36030>
- Sims, J. R., Zimmer, J. A., Evans, C. D., Lu, M., Ardayfio, P., Sparks, J., Wessels, A. M., Shcherbinin, S., Wang, H., Monkul Nery, E. S., Collins, E. C., Solomon, P., Salloway, S., Apostolova, L. G., Hansson, O., Ritchie, C., Brooks, D. A., Mintun, M., Skovronsky, D. M., & TRAILBLAZER-ALZ 2 Investigators. (2023). Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial. *JAMA*, 330(6). <https://doi.org/10.1001/jama.2023.13239>
- Solimando, L., Fasulo, M., Cavallero, S., Veronese, N., Smith, L., Vernuccio, L., Bolzetta, F., Dominguez, L. J., & Barbagallo, M. (2022). Suicide risk in caregivers of people with dementia: a systematic review and meta-analysis. *Aging Clinical and Experimental Research*. <https://doi.org/10.1007/s40520-022-02160-6>
- Sperling, R. A., Donohue, M. C., Raman, R., Sun, C.-K., Yaari, R., Holdridge, K., Siemers, E., Johnson, K. A., & Aisen, P. S. (2020). Association of Factors With Elevated Amyloid Burden in Clinically Normal Older Individuals. *JAMA Neurology*, 77(6), 735. <https://doi.org/10.1001/jamaneurol.2020.0387>
- Stein Emil Vollset, Ababneh, H. S., Yohannes Habtegiorgis Abate, Cristiana Abbafati, Rouzbeh Abbasgholizadeh, Mohammadreza Abbasian, Hedayat Abbastabar, Magied, A., Samar Abd ElHafeez, Abdelkader, A., Abdelmasseh, M., Sherief Abd-Elsalam, Abdi, P., Abdollahi, M., Meriem Abdoun, Abdullahi, A., Abebe, M., Olumide Abiodun, Richard Gyan Aboagye, & Abolhassani, H. (2024). Burden of disease scenarios for 204 countries and territories, 2022–2050: a forecasting analysis for the Global Burden of Disease Study 2021. *Lancet*, 403(10440), 2204–2256. [https://doi.org/10.1016/s0140-6736\(24\)00685-8](https://doi.org/10.1016/s0140-6736(24)00685-8)
- United Nations. (2019). Global Issues: Ageing. United Nations. <https://www.un.org/en/global-issues/ageing>
- U.S. Food and Drug Administration. (2021, June 7). FDA Grants Accelerated Approval for Alzheimer's Drug. FDA. <https://www.fda.gov/news-events/press-announcements/fda-grants-accelerated-approval-alzheimers-drug>
- Van der Mussele, S., Le Bastard, N., Saeuens, J., Somers, N., Mariën, P., Goeman, J., De Deyn, P. P., & Engelborghs, S. (2014). Agitation-associated behavioral symptoms in mild cognitive impairment and Alzheimer's dementia. *Aging & Mental Health*, 19(3), 247–257. <https://doi.org/10.1080/13607863.2014.924900>

- Van Dyck, C. H., Swanson, C. J., Aisen, P., Bateman, R. J., Chen, C., Gee, M., Kanekiyo, M., Li, D., Reyderman, L., Cohen, S., Froelich, L., Katayama, S., Sabbagh, M., Vellas, B., Watson, D., Dhadda, S., Irizarry, M., Kramer, L. D., & Iwatsubo, T. (2022). Lecanemab in Early Alzheimer's Disease. *New England Journal of Medicine*, 388(1), 9–21. <https://doi.org/10.1056/nejmoa2212948>
- Vermunt, L. (2019). Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. *Alzheimer's & Dementia*, 15(7), 888–898. <https://doi.org/10.1016/j.jalz.2019.04.001>
- Vos, S. J., Xiong, C., Visser, P. J., Jasielec, M. S., Hassenstab, J., Grant, E. A., Cairns, N. J., Morris, J. C., Holtzman, D. M., & Fagan, A. M. (2013). Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. *The Lancet Neurology*, 12(10), 957–965. [https://doi.org/10.1016/S1474-4422\(13\)70194-7](https://doi.org/10.1016/S1474-4422(13)70194-7)
- Wegmann, S., Biernat, J., & Mandelkow, E. (2021). A current view on Tau protein phosphorylation in Alzheimer's disease. *Current Opinion in Neurobiology*, 69, 131–138. <https://doi.org/10.1016/j.conb.2021.03.003>
- World Health Organization. (2023, March 15). Dementia. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/dementia>
- World Health Organization. (2013). Definition and list of health professionals. Transforming and Scaling up Health Professionals' Education and Training - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK298950>

Causes of Non-Verbal Learning Disorder and the Need for Inclusion in the DSM-5-TR

By Advaita Vemuri and Zahara Khan

Abstract

Nonverbal Learning Disorder is a neurodevelopmental disorder initially hypothesized by Dr. Byron P. Rourke in 1989, since emerging as the subject of numerous debates regarding its validity and whether it should be considered an official diagnosis in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR). This paper seeks to consolidate current information regarding NVLD in order to assert that the DSM-5-TR must officially include NVLD within the neurodevelopmental disorder category. Comparisons between NVLD and existing categories of neurodevelopmental disorders serve to cement its place within the category of neurodevelopmental disorders while highlighting the unique characteristics of NVLD, demonstrating the necessity of an official diagnosis. Through extensive research regarding what is currently known about NVLD and other neurodevelopmental disorders, we have isolated the cause of the symptoms of NVLD to deficiencies in the white matter of the right hemisphere of the cerebral cortex and smaller sized splenia in the corpus callosum, identifying a possible connection between these two phenomena and the occurrence of NVLD's symptoms, validating the need for further research into the causes of NVLD. In assembling a complete psychological profile, deficits were identified in five main categories: mathematical reasoning, visuospatial awareness, higher-order comprehension, social communication, and executive functions. Despite the magnitude of officializing an entirely new disorder, biological and behavioral disruptions unique to NVLD validate its existence as a distinct neurodevelopmental disorder and warrant its addition to the DSM-5-TR. Additional research is vital for assembling a proper neurological and symptomatic profile, cause, and treatment.

Introduction

Nonverbal Learning Disorder, commonly abbreviated as NLD or NVLD, is a neurodevelopmental disorder that is not yet officially recognized in the DSM-5-TR. The condition manifests in a variety of symptoms that impact brain function, which include visuospatial and motor skill deficiencies that impair daily life. These symptoms are known to emerge during childhood, and remain present in a person throughout their adulthood. Despite its recognition by many members of the field of psychology, Nonverbal Learning Disorder is currently not included as an official diagnosis. However, due to the prevalence of the disorder and significant issues that people with NVLD face, Nonverbal Learning Disorder should be added to the DSM-5-TR. Currently, there is a dearth of information regarding subjects such as how to properly treat those with NVLD and the root biological cause of the disorder. This makes providing the proper support to children who struggle with the unique set of deficiencies NVLD involves extraordinarily difficult. Further research into treatments for NVLD is crucial for confirmation of the physiological cause and possible ways to treat or alleviate symptoms, which

is why the DSM-5-TR needs to recognize it as a distinct disorder and grant it the recognition it needs.

Background

Nonverbal Learning Disorder (NVLD)

Nonverbal learning disorder is a disorder specified by a variety of issues regarding visuospatial skills, higher-order and abstract comprehension, and executive functioning. NVLD is classified as a neurodevelopmental disorder because it is the result of neurological deficiencies that manifest during childhood and adolescence that last throughout adulthood, and therefore classified as a mental disorder as it is a significant disturbance that impairs important cognitive functions on a day to day basis. The cognitive impairments associated with NVLD encompass a myriad of symptoms that impact the quality of life for people with the disorder, and their ability to function effectively on a daily basis. Nonverbal learning disorder was first identified in Johnson and Myklebust's 1967 book *Learning Disabilities: Educational Principles and Practices*. However, the modern day understanding of NVLD was due to Dr. Byron P. Rourke's work in identifying a comprehensive model for NVLD's symptoms. Rourke's works, *Syndrome of Nonverbal Learning Disabilities: The Syndrome and the Model* (1989) and *Syndrome of Nonverbal Learning Disabilities: Neurodevelopmental Manifestations* (1995), have informed most of the modern discussion regarding NVLD. In his books, Rourke developed his hypothesized symptomatic profile for NVLD and presented a white matter model for the condition, in which he proposed that the depicted symptoms of NVLD were due to deficiencies located in the white matter of the right hemisphere of the brain. Ever since then, NVLD has gained recognition from members of the scientific field and select members of the public as more people investigated and performed studies regarding the condition. However, even though NVLD has been known of as a psychological condition for over half a century, it is still relatively lacking in information. There is no official symptomatic profile or biological cause for NVLD currently, and no disorder specific treatment or support plan for children with NVLD. Most notably, despite the recognition NVLD has amassed over time, it has not yet been recognized in the DSM-5-TR as an official, legitimate diagnosis under the neurodevelopmental umbrella. There is much debate over whether this addition should even occur, but research done over the course of the decades since NVLD has become known has validated its existence as a disorder.

Categories of Neurodevelopmental Disorders in the DSM-5-TR

A neurodevelopmental disorder is a condition that affects how the brain functions, which ranges from mild to severe impairments. These disorders are known to manifest during childhood and adolescence, resulting in a variety of deficiencies in multiple areas of life for the afflicted child that typically do not go away as they age. They stem from issues regarding the development of the nervous system during childhood, in both the brain and spinal cord, and

symptoms may continue well into adulthood, being a lifelong impairment. There are currently six main categories of neurodevelopmental disorders in the DSM-5-TR: Autism Spectrum Disorder, Attention-Deficit/Hyperactivity Disorder, Intellectual Development Disorders, Communication Disorders, Motor Disorders, and Specific Learning Disorders. Nonverbal Learning Disorder is not currently recognized as a disorder among these disorders, despite displaying the classifying conditions for a neurodevelopmental disorder and remaining distinct from existing categories. Autism Spectrum Disorder, commonly known as Autism, is characterized by its spectrum of behavioral symptoms such as impairment of social skills, communication problems, repetitive actions, and circumscribed interests. It is the result of genetic inheritance, as well as several biological factors within the brain. A larger hippocampus, smaller amygdala with increased activation, and frontal lobe abnormalities within the prefrontal cortex are associated consistently with Autism Spectrum Disorder and assumed to be causation factors (Rylaarsdam & Guemez-Gamboa 2019). Attention-Deficit/Hyperactivity Disorder, more commonly referred to as ADHD, is characterized by sporadic periods of mild to extreme inactivity, hyperactivity, and impulsivity. Biologically, ADHD is characterized by a decrease in dopaminergic receptors in the frontal lobe. A smaller anterior cingulate gyrus and dorsolateral prefrontal cortex are correlated with ADHD, as well as decreased activity in the frontostriatal region of the brain (Magnus et al. 2023). Intellectual Development Disorders are caused by a variety of different factors, such as chromosomal conditions (ex: Down Syndrome), environmental components, metabolic conditions (ex: increased bilirubin levels), infections, and potentially unexplained reasons. Behavioral symptoms of this disorder include walking later than peers, difficulties with talking, remembering, and understanding social cues, and issues with problem solving that impair day to day activity (U.S. National Library of Medicine 2013). Communication Disorders, which include both speech and language disorders, are caused by possible toxin exposure during fetal development, brain abnormalities, or a possible genetic factor (The Children's Hospital of Philadelphia 2014). It is characterized by communication difficulties, as well as repetition and addition of sounds to speech (American Psychiatric Association 2022). Specific Learning Disorders, characterized by difficulties in reading and language processing, are theorized to be caused by parietal and premotor cortex deformities and genes associated with neural development (Saini et al. 2023). And lastly, Motor Disorders, which are uncontrollable movements or actions of the body, like tics or spasms, that are theorized to be caused by genetic factors and environmental factors, like exposure to alcohol while in fetal development (U.S. National Library of Medicine 2013).

Table 1.

Comparison of currently included neurodevelopmental disorders in the DSM-5-TR.

Disorder	Behavioral Symptoms	Biological Causes
Autism Spectrum Disorder	- Impairment of social skills - Communication problems - Repetitive actions - Circumscribed interests (Rylaarsdam & Guemez-Gamboa 2019)	- Genetic inheritance - Larger hippocampus - Smaller amygdala with increased activation - Frontal lobe abnormalities within the prefrontal cortex (Rylaarsdam & Guemez-Gamboa 2019)
Attention Deficit/Hyperactive Disorder (ADHD)	- Sporadic periods of: - Mild to extreme inactivity - Hyperactivity impulsivity (Magnus et al. 2023)	- A decrease in dopaminergic receptors in the frontal lobe - A smaller anterior cingulate gyrus and dorsolateral prefrontal cortex - Decreased activity in the frontostriatal region of the brain (Magnus et al. 2023)
Intellectual Development Disorders	- Walking later than peers - Difficulties with talking, remembering, and understanding social cues - Issues with problem solving that impair day to day activity (U.S. National Library of Medicine 2013)	- Chromosomal conditions (ex: Down Syndrome) - Environmental components - Metabolic conditions (ex: Increased bilirubin levels) - Infections - Potentially unexplained reasons (U.S. National Library of Medicine 2013)
Communication Disorders	- Communication difficulties - Repetition and addition of sounds to speech (American Psychiatric Association 2022)	- Possible toxin exposure during fetal development - Brain abnormalities - A possible genetic factor (The Children's Hospital of Philadelphia 2014)
Specific Learning Disorders	- Difficulties in reading and language processing (Saini et al. 2023)	- Parietal and premotor cortex deformities - Genes associated with neural development (Saini et al. 2023)
Motor Disorders	- Uncontrollable movements or actions of the body (U.S. National Library of Medicine 2013)	- Genetic factors and environmental factors - Exposure to alcohol while in fetal development (U.S. National Library of Medicine 2013)

Controversy with Recognizing NVLD as a Diagnosis

NVLD's validity as a disorder has been contested ever since its original conception in the 1980s due to concerns regarding its similarities with other neurodevelopmental disorders and lack of a defined symptomatic profile. It has oftentimes been compared to Autism Spectrum

Disorder (ASD) and Attention Deficit/Hyperactivity Disorder (ADHD) in regards to its symptomatic and biological profile, resulting in assertions that NVLD does not warrant its own category in the DSM-5-TR and is instead a subcategory of Autism Spectrum Disorder or ADHD. Due to individuals with NVLD struggling with social situations and perceiving nonverbal cues, they are often misconceptualized as having Autism Spectrum Disorder. Individuals with NVLD often struggle with focusing and organizing under a big picture, which can be misconstrued by professionals as symptoms of ADHD. Additionally, NVLD is not particularly well-known, meaning that research on the subject has been rather limited in its scope. However, existing research presents potential biological causes that are distinct from other neurodevelopmental orders, distinguishing NVLD as its own category. Moreover, NVLD's behavioral profile revolves around different core traits than Autism Spectrum Disorder, or any other neurodevelopmental disorder. Arguments that seek to classify NVLD as a subset of existing neurological disorders overlook key traits of the disorder, instead categorizing them as attentional issues or as a similar disorder, which comes at the cost of misunderstanding its symptoms and its severity (Burkhardt 2019). Attentional issues are much more surface level, and with time it can be determined if the issue is deeper. Disorders like Autism Spectrum Disorder and ADHD do have areas of overlap with NVLD, which makes it difficult to specify what disorder to attribute symptoms to. However, while the symptomatic profiles of these disorders may appear similar, the deficiencies children with NVLD possess are in actuality based on different symptoms from the other aforementioned neurodevelopmental disorders, with these traits being severe enough to be considered a different disorder altogether.

NVLD'S Unique Behavioral Expression

NVLD is characterized by a variety of neurological deficiencies that are expressed in different ways depending on the individual. Individuals with NVLD can be said to have five main deficiencies: visuospatial awareness, higher-order comprehension, social communication, math concepts, and executive functions (Miller). Visuospatial understanding is perhaps NVLD's most distinctive deficit, as it is central to the holistic symptomatic profile. Visuospatial processing refers to the ability to process the location of an object in space through sight, and all of the mental processes this entails. Cognitive abilities such as the creation of visual representations and manipulated mental representations fall within the boundaries of visuospatial processing (Possin 2010). This disruption in visuospatial understanding contributes to the issues with mathematics that children with NVLD face, as math tends to be a predominantly spatial subject in school. Carleen Franz noted how "their work is characterized by various sizes of numbers, inaccurate vertical alignment, and slanted horizontal positioning", which demonstrates how lack of spatial awareness serves as a detriment to that mathematical abilities of someone with NVLD (2000). This leads to children with NVLD struggling in school and seeing their performance suffer due to a lack of proper support for their neurological needs. Another notable aspect of NVLD is the difficulties that arise in regards to higher-order comprehension, which results in learning difficulties associated with abstract concepts. High-order thinking requires a

level of cognitive ability that goes beyond mere memorization, and requires skills such as visualization, abstract models of thought, and application of knowledge. Individuals with NVLD tend to do well with rote memorization, but lack the capacity to engage with the higher-level mental processes that are associated with later grades. This further contributes to difficulties with mathematics and visual learning. Mathematical concepts become more and more pertinent and abstract as children progress through school and reach upper grade levels, which tend to require a level of comprehension that children with NVLD struggle with. These struggles may stem from the inability to engage with mental processes and abstract representations due to visuospatial deficiencies. In contrast to this difficulty with math, children with NVLD tend to have much higher scores for subjects such as reading, demonstrating how NVLD's effects manifest in different ways for different mental processes.. Visual learning involves grasping overarching concepts and drawing conclusions based on what is implied, which is a major difficulty for individuals with NVLD, who tend to have issues with processing details and viewing a main idea or theme that isn't clearly stated. Another main characteristic, difficulty engaging in social communication, is a trait shared by numerous other neurodevelopmental disorders, such as Intellectual Development Disorders and Autism Spectrum Disorder. In regards to NVLD, it is likely that people with the disorder struggle to pick up on nonverbal social cues, thus negatively impacting their social communication skills (Franz 2000). Such limitations can result in isolation from peers or a sense of inadequacy in regards to navigating social situations, which can be harmful to a child's mental health as they grow up. It is important to identify what exactly is the cause of a child's struggle with socialization so as to best provide the support they need to connect effectively with their peers in their young, malleable age. Due to the disruptions individuals with NVLD face in regards to processing abstract visuals, it is possible that providing them with straightforward verbal explanations can help with effective processing of information. Furthermore, like with most neurodevelopmental disorders, executive dysfunction is a prominent issue in NVLD as well, even serving as a precursor to several specific neurodevelopmental disorders (Otterman et al. 2019). Executive functioning refers to a person's capacity for exercising their working memory, organization, focus, cognitive flexibility, and self-control (Diamond 2013). These sorts of abilities are necessary for a multitude of areas of life, especially in regards to academic success. Due to the issues children with NVLD face in regards to higher-level comprehension, they face difficulties that impact their executive functioning skills, which can serve as a detriment to both their school and daily life. This lack of ability to effectively carry out these executive functions is known as executive dysfunction. Research has found that children with NVLD are weaker in every domain of executive functioning when compared to neurotypical children, which intensifies the other daily struggles associated with the disorder (Kouhbanani et al. 2020). NVLD impacts multiple aspects of life, ranging from academics to social navigation and daily functioning. Varying in severity, like other neurodevelopmental disorders, it is a lasting detriment that becomes increasingly difficult to deal with without help or aid as time goes on and issues are exacerbated. The combination of struggles that people with this disorder face uniquely center around visuospatial and abstract

comprehension, which warrants a separate category in the DSM-5-TR in order to best assist them.

Hypothesis for the Biological Cause of NVLD

There is not yet an approved, concrete theory for the biological causes of NVLD in individuals. It is a condition that lacks important research in regards to determining causation and treatments, with few studies due to its lack of an official status. But while there may not be an official theory regarding biological causes of this disorder, what research there is has provided a general idea. Children with NVLD have been observed to have a number of differences between their brains and a regular brain, particularly through MRI scans. One notable observation is how those with NVLD have smaller spleniums, which is a part of the corpus callosum (Fine et al. 2013).

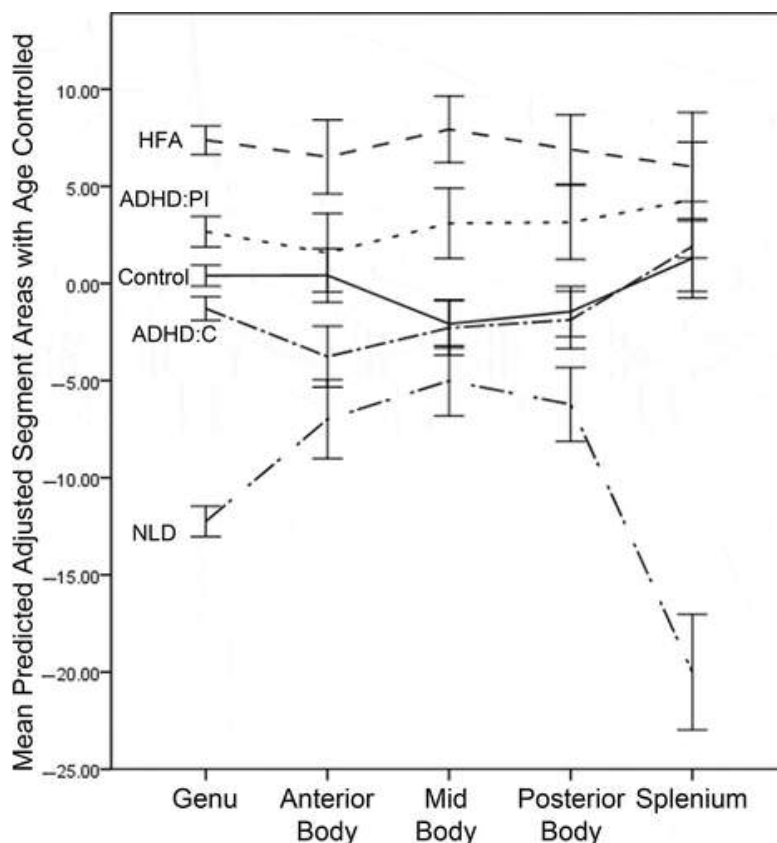


Fig 1: Approximate area of corpus callosum components by group. The table depicts the approximate area of the five main components of the corpus callosum. The mean area for the splenium is notably much lower for Nonverbal Learning Disorder than for any other group, with error bars up to 95% confidence. (Fine et al. 2013)

The corpus callosum, made up of the genu, anterior body, mid body, posterior body, and splenium, is a vital part of the brain, as it connects the left and right cranial hemispheres. The splenium plays a role in transferring information between the two hemispheres of the brain, and is involved in visuospatial processing, language, and calculation scores (Blaauw & Meiners,

2020). This matches well with the identified deficiencies of NVLD, particularly the difficulties with visuospatial processes and mathematical processes. This anatomical difference is unique to NVLD, and is not detected in a normal brain, or in the brain of someone with a different neurodevelopmental disorder, such as ADHD or Autism Spectrum Disorder. The presence of a smaller splenium is not the only noted difference, however; it is noted by Dr. Byron P. Rourke, the proponent of the most recognized model of NVLD, that people with NVLD tend to show signs of deficiencies in the white matter of the right hemisphere of their cerebral cortex. Rourke had in fact presented and supported a white matter model of NVLD. He had found that the symptomatic profile of NVLD showed signs of correlation with the presence of damage to the white matter of the right hemisphere of the brain (Rourke 1987). The right hemisphere is responsible for a variety of functions, including spatial thinking and image processing, while damage to this hemisphere can result in difficulties with nonverbal memory (Maldonado & Alsayouri 2023). This is consistent with the symptomatic profile of NVLD, which is characterized by its deficiencies with nonverbal cues and visuospatial awareness. Additionally, research has found that certain mathematical processes are localized to the right hemisphere of the brain, such as the “operation-specific spatial layout required for complex arithmetical procedures” (Salillas et al. 2023). Thus, it can be inferred that the visuospatial disruption associated with NVLD is what contributes to the difficulties with mathematics, and as these functions are linked to the right hemisphere, deficiencies in the white matter could be the source of these struggles. However, Rourke had later noted that NVLD symptoms were almost always demonstrated in case of agenesis of the corpus callosum (Rourke et al. 2002). An agenesis of the corpus callosum entails a missing or partially missing or damaged corpus callosum, which is consistent with the idea that the splenium plays a dominant role in NVLD’s occurrence. The White Matter Model and the splenium model can be linked however, as the corpus callosum itself is composed of white matter. Since Rourke’s White Matter Model suggests that damage or deficiencies to the myelinated axons that make up white matter are the cause of NVLD, it is worth investigating whether the reduced splenium size contributes to these white matter deficits. This is furthered by results indicating that there are more splenial connections from the right to the left hemisphere than the left to the right hemisphere (Putnam et al. 2010). This hemispheric asymmetry may explain the connection between splenial size and deficiencies in the white matter of the cerebrum’s right hemisphere. Furthermore, Putnam and his team’s results indicated the possibility that the hemispheric asymmetry in regards to splenial connections connected to the right hemisphere’s “superiority for higher order perceptual processing” (Putnam et al. 2010). This creates a connection between the white matter deficiencies associated with NVLD and the struggle with both visuospatial processing and higher-order comprehension. Further testing and research into NVLD is necessary in order to confirm this as a biological cause. Methods such as diffusion tensor imaging have been suggested as a means of determining the impact a smaller splenium size would have on its connections with other white matter connections in the posterior of the brain (Fine et al. 2013). This method of imaging examines the Brownian motion of the water particles within the cerebrum, and records this motion in order to develop a map of the

neural circuits of the brain (Ranzenberger & Snyder 2021). Diffusion tensor imaging focuses on tracking white matter, and so could demonstrate the impacts the reduction in splenium size has on the physiological connections of the brain. Upon further analysis, various aspects of the symptomatic profile of NVLD can thus be assigned to any observed deficiencies in the white matter. These results could be furthered by a functional MRI scan, which tracks the flow of oxygenated blood within the brain. Because the flow of oxygenated blood shows what regions of the brain are involved in certain mental processes, comparing the fMRI of a person with NVLD with a neurotypical person can demonstrate how brain activity is impacted by a smaller splenium and deficiencies in the white matter of the right hemisphere.

Case Study

In order to determine if the symptoms of NVLD can be attributed to a connection between a smaller splenium and white matter deficiencies, a case study of an acallosal individual can be examined. This case study was carried out by Patrick Panos, Scott S. Porter, Angelea Panos, Richard Gaines, and Philip Erdberg, and all of the data and associated conclusions are attributed to them and their team. Patrick Panos is currently an associate professor of social work at the University of Utah, with a PhD in counseling psychology from Brigham Young University and a BA in psychology from the University of Utah. Scott S. Porter is a clinical neuropsychologist at Saible Neuropsychology, with a PhD in clinical neuropsychology/behavioral neuropsychology from Brigham Young University and a BA in psychology from the University of Utah. Angelea Panos is the Substance Use Disorder Counseling program director and professor in the psychology department at Utah Valley University. Philip Erdberg was The subject of the case study is an 11 year-old white male who is referred to as “K”. K had previously been diagnosed with disorders such as Attention Deficit/Hyperactive Disorder and had a history of hyperactivity and aggression. It was found that he was born acallosal, meaning his corpus callosum was almost entirely absent. A series of tests were carried out to determine his lateral abilities, with K being found to have deficits in most of the same areas as predicted by Rourke’s model of NVLD. For example, K had demonstrated difficulty with “integrative tasks such as visual–spatial organization” (Panos et al. 2001). However, the team came to the conclusion that the severity of K’s symptoms overshoot the predictions of Rourke’s model, as his struggles with skills such as fine-motor coordination were far greater than expected. K also was found by the team to have deficits with language and speech that were not part of the NVLD model. Thus, the team concluded that Rourke’s model was not entirely conclusive about the struggles that may arise as a result of white matter deficiencies, and that continued testing of acallosal children was necessary. All of the previously mentioned results are credited to the aforementioned researchers who carried out the case study. Based on the results, it is likely that a link between the corpus callosum, white matter of the brain, and the symptoms of NVLD are existent. Since the team found K’s symptoms to be much greater than expected, it is possible that a less severe impairment of the corpus callosum is the cause of symptoms of NVLD that would align with Rourke’s model in terms of severity. Thus, a

smaller or deficient splenium, a part of the corpus callosum, could be the cause of NVLD's general symptomatic profile. It has already been surmised that the splenium bears a connection with the right hemisphere of the brain, specifically in regards to its white matter. This can be further determined when examining the MRI scan of K's brain taken by the team, which was compared with a regular brain.

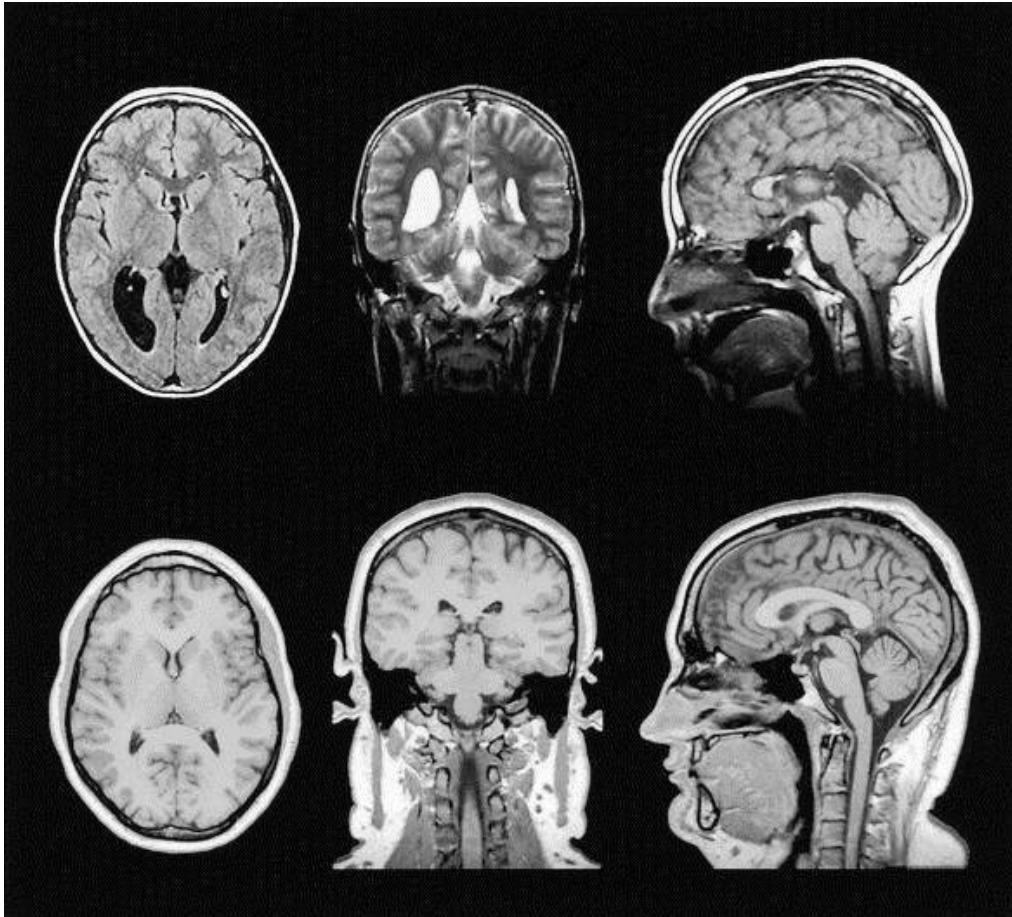


Fig 2: Comparison of K's MRI scans with standard MRI scans. The top row depicts MRI scans for K's brain from three different angles, while the bottom row depicts a standard MRI scan. The imaging in K's brain uses white to depict ventricular space (Panos et al. 2001).

The team had concluded from this data that K's brain had a larger ventricle on the right cranial hemisphere, contributing to the idea that splenial deficiencies are localized asymmetrically. A mostly complete lack of the corpus callosum, which is entirely composed of white matter, led to severe symptoms that were associated with NVLD. However, it also presented symptoms that were not included in Rourke's model at all because much more than just the splenium was missing. So, confining the biological cause of NVLD to the splenium can narrow the symptoms to just those associated with NVLD, as well as match the predicted severity of deficiencies associated with the condition. This can be determined through the aforementioned imaging studies of the brains of people with NVLD or a recorded smaller or nonexistent splenium in order to track the white matter deficiencies, as well as ascertaining how

a splenial defect could affect cognitive processes. This methodology would present the biological basis of NVLD, and facilitate the study of how the physiological differences of a brain with a smaller or deficient splenium connect with deficiencies in the white matter of the right hemisphere and ascertain how this results in the symptoms of NVLD.

Differences Between NVLD and Official Neurodevelopmental Disorder Categories

NVLD's status as a neurodevelopmental disorder causes it to bear similarities with other disorders that fall under the same umbrella. These similarities cement the place of NVLD within the category of neurodevelopmental disorders, while significant enough differences present NVLD as a separate disorder. The issues children with NVLD face regarding higher-order reasoning and social communication often are misconstrued as being symptoms of Autism Spectrum Disorder. ADHD is also a common misdiagnosis due to the difficulties with school that children with NVLD tend to experience. However, these diagnoses fail to account for the symptoms that NVLD centers around, mainly being issues with mathematical concepts, visuospatial deficiencies, and higher-order thinking processes. People with other neurodevelopmental disorders may experience symptoms related to some of these main categories, but NVLD's unique presentation of these main symptoms qualify it to be considered a separate disorder in the DSM-5-TR. Its core traits are different from other neurodevelopmental disorders, and any attempts to classify NVLD as under a preexisting condition would fail to account for how children are affected by the particular characteristics of NVLD. For example, children with NVLD are shown to have significant differences in regards to their pragmatic language skills with children that have Autism Spectrum Disorder. Children with NVLD had much better pragmatic language skills than children with Autism Spectrum Disorder, demonstrating how the symptoms of NVLD resulted in behaviors and struggles that were distinct from those associated with Autism (Mammarella et al. 2022). However, these results would likely have been very different if the comparison was made in regards to mathematics, as difficulties with mathematics is one of the primary ways that NVLD manifests. Such struggles are not associated with the symptomatic profile of Autism Spectrum Disorder, demonstrating how the two conditions have vastly different impacts. Thus, it can be concluded that NVLD is its own unique category of disorder, and ought to be treated as such in the DSM-5-TR.

Diagnostic Criteria

NVLD can be said to have a few main traits that predominantly make up its symptomatic profile. The primary deficiency associated with NVLD is visuospatial in nature, and can be said to be a key symptom of the disorder. Katherine Possine, a doctor with a PhD in clinical psychology and a focus on neurodegenerative diseases, elaborated upon how this can be applied to a diagnostic criteria, with traits such as a low visuospatial intelligence, poor visuospatial memory, and a discrepancy in verbal and visuospatial intelligence serving as prime characteristics of the disorder (Possin 2010). If a child is struggling with spatial orientation, facing issues with identifying where an object is located based on visual cues, or is

demonstrating similar forms of spatial disorientation in relation to sight and vision, they can be identified as presenting the necessary symptoms associated with this particular deficiency. Additionally, having difficulties with mathematics in school is a good indicator of NVLD, particularly if the student does better with reading than with arithmetic and other mathematical concepts. In line with struggles regarding school, someone who appears to do well with rote memorization but has trouble with more abstract thought patterns and conceptual application would also demonstrate symptoms consistent with NVLD. This can manifest in a variety of ways in school and later on in life, but the discrepancy between academic math and reading and poor performance regarding mathematics are the most prominent examples. In general, struggles regarding higher-order comprehension beyond stated facts and memorization is a prime characteristic of the disorder, and is indicative of a possibility that the child has NVLD. Struggles regarding socialization and fine motor skills can also serve as part of the diagnostic criteria for NVLD, particularly if these issues stem from difficulties with processing nonverbal cues and understanding spatial orientation. As is common with official neurodevelopmental disorders, executive dysfunction is another major symptom of NVLD, and can indicate a deficiency regarding the formation of mental models and abstract comprehension. Ultimately, the criteria for diagnosing NVLD is rooted in isolating a cognitive visuospatial deficiency that results in difficulties with higher-order comprehension skills, mathematical reasoning, abstract thought patterns, social skills, and fine motor skills.

Prevalence

While NVLD has become much more well known over the years, its lack of recognition as an official diagnosis most likely has impacted statistics regarding its prevalence in the population. Current data samples have suggested a prevalence of 3% to 4%, meaning that 3-4% of the adolescent population demonstrate symptoms of and have NVLD (Margolis et al. 2020). When multiplied against the more than 300 million people living in just the United States (United States Census Bureau 2024), this percent becomes around 10 million people. This is a significant portion of the population, representing millions of children and adults alike struggling with NVLD and most likely receiving little to no proper support due to the lack of a proper diagnosis. Moreover, it is possible that in actuality, the actual prevalence of the disorder is higher than was previously thought. Information regarding the presentation of NVLD tends to center around children, neglecting adults that would have been impacted into their adulthood. Also, other factors such as familial links or environmental risks are not currently tested for. There has been no public research up until this point on the effects of any certain demographic, such as racial, social, or economic status, playing a role in the occurrence rate of NVLD. Thus, further studies must be done in order to determine if a significant portion of the population has NVLD, and what percentage of certain populations has the disorder.

Comorbidity

Information regarding comorbidity of NVLD is extraordinarily difficult to reliably assess at the moment, as a lack of official recognition for NVLD has contributed to major inconsistencies regarding its diagnosis. Thus, the current data for who has and doesn't have NVLD may have a large deviation from the actual population of people with the disorder, making comorbidity a difficult thing to track. However, it can be reasonably surmised that disorders such as Autism Spectrum Disorder and Attention Deficit/Hyperactive Disorder have a tendency to occur alongside NVLD, due to behavioral symptom similarities. For example, individuals with Autism Spectrum Disorder and individuals with NVLD have been observed to be more literal about words and phrasing than neurotypical people, as well as being more oblivious with recognition of social cues. However, the similarities end when abstract concepts are brought in. A major trait of NVLD is its difficulties with abstract ideals, especially in regards to mathematics in an academic context. Individuals with Autism Spectrum Disorder can still perform well with mathematics and other abstract concepts encountered in schools and daily life, and many excel in them. Additional co-occurring symptoms include depression and anxiety, which is common with neurodevelopmental disorders (Mammarella et al. 2022). The comorbidity of NVLD must continue to be tracked in order to ensure an accurate understanding of what symptoms and disorders may occur alongside NVLD, and the commonality of these comorbidities. However, as there are behavioral and assumed biological differences in presentation between NVLD and any potential co-occurring symptoms and conditions, NVLD can be considered to be unique enough to constitute a separate diagnosis category in the DSM-5-TR.

Why the DSM-5-TR Must Include NVLD

While there is a general idea of what causes NVLD, there is still no definitive theory regarding the biological component. Further research into the exact biological cause of NVLD would allow psychologists to understand the behavioral nuances of people with NVLD and why they occur. It should also be noted that nearly all current research into NVLD focuses on youths, mainly elementary school aged children. There is a major lack of information about what NVLD looks like in an adult, and how its symptoms continue to manifest as time goes on. NVLD, following the trends of other neurodevelopmental disorders, shows no signs of disappearing as affected children become adults. Future studies about NVLD should include adults with the disorder in order to provide a broader perspective on how the magnitude of NVLD's impact on an individual. There is also the matter of treatment for those with NVLD. Many individuals with NVLD end up getting diagnosed solely with ADHD or Autism Spectrum Disorder due to the lack of a definitive comprehensive diagnostic criteria. This results in people with NVLD not receiving proper support, such as being prescribed medications that are meant for treating other neurodevelopmental disorders such as ADHD. Due to the biological differences between NVLD and ADHD, these prescriptions have no effect on mitigating the effects of NVLD, leaving the individual without actual support. Additionally, an ADHD or Autism Spectrum Disorder diagnosis does not draw enough attention to the issues children with NVLD face regarding

academic mathematics and visuospatial awareness (Miller). In order to provide proper support and treatment to people with NVLD, it must be recognized as a legitimate diagnosis and receive additional research into causes and ways to address it, as there is a lack of both awareness and incentive about NVLD. Not many know about the condition, since it is not an official one, and there is little to no incentive to do research on a condition that does not exist within the DSM-5-TR. There are currently no medications or official treatment plans for NVLD nor are there any tests with reliable data regarding the effect of medications on those with NVLD, which limits current plans of action for assisting individuals with the disorder. However, steps can be taken to accommodate for the limitations NVLD presents in classrooms. Sue Thompson's handbook (1997) specifies how children can be given direct verbal strategies and modified expectations for their performance in the classroom in order to account for these weaknesses. A more modern example of educational support is demonstrated by Winston Preparatory School, as they present a website page detailing the differences between Autism Spectrum Disorder and NVLD alongside their commitment to differentiating between the two and providing specific guidance and assistance to their students (Winston Preparatory School). Additional forms of support for abstract mental processes can also be beneficial towards children with NVLD and account for difficulties with formulating mental models. Such accommodations can greatly reduce the strain of getting through school for children with NVLD and provide much needed support for their difficulties. Moreover, different forms of training and therapy can help the child build up their cognitive processes in domains that are weak. Perceptual-motor training involves various motor activities and exercises that are designed to work the cardiovascular system and improve cognitive function over time through coordination and engagement. Such activities have been shown to improve the executive function of children with NVLD in particular, suggesting that perceptual-motor training or a similar approach can be beneficial towards improving the daily function of those with the disorder (Kouhbanani et al. 2020). Because of how deeply executive functioning is intertwined with the other main areas of the disorder, such as visuospatial and abstract processing, perceptual-motor training can contribute to a better overall quality of life for individuals with NVLD. In order to verify this, further testing and analysis of the impacts of perceptual-motor training on people with NVLD must be ascertained with a holistic focus. Using people with a typical neurological development as the control group, an experimental group of people with NVLD can have their cognitive abilities assessed in all areas affected by NVLD and weighed against the control group. The experimental group will be divided by age range, and can receive perceptual-motion training over time, with consistent assessments to examine any change in their cognitive abilities. The neurotypical group will be similarly categorized by age ranges, and each age range of the control group will be respectively compared with the experimental group. This basic experimental procedure could provide a possible basis for further treatment options by analyzing what symptoms are alleviated and what areas still need to be addressed, as well as determining how NVLD affects people at a large range of different ages. There is no doubt that further general testing is necessary for determining the most efficient way to support individuals with NVLD. NVLD is a serious disorder that

uniquely impacts every aspect of an individual's life, qualifying it as an entry in the DSM-5-TR. With this addition, NVLD can receive the attention necessary for further research and analysis of everything related to it.

Conclusion

Adding a new disorder to the DSM-5-TR is no small task: it will come with many changes, both good and bad. However, it needs to be done in the case of NVLD. It has been a topic of controversy since it was introduced by Rourke, with several critiques regarding its validity and similarity with other official disorders such as the aforementioned Autism Spectrum Disorder and ADHD. Performance deficits are also oftentimes perceived as motivational or attentional issues. However, differences in behavioral symptoms and consistent findings in reports prove that NVLD does have the necessary qualifications to be included as a disorder within the DSM-5-TR. Introducing it as an official disorder will create a new issue: It will cause several individuals with diagnosed Autism, ADHD or other conditions with similarities to NVLD to need to be retested and potentially rediagnosed due to the addition of a new disorder. NVLD is different in behavioral symptoms and biological factors in comparison with similar conditions, but is similar enough to be misdiagnosed very frequently, which will warrant much retesting. It will be an arduous task to accomplish, but one that needs to be done in order for individuals struggling with NVLD to get the help they need. Treatment for similar disorders does not always work, especially with medications, and with the large number of possible NVLD cases in the world, the need for specialized treatment is immense. Due to its severity, prevalence, and qualifications, NVLD needs to be recognized as an official diagnosis in the DSM-5-TR, since the needs outweigh the negative repercussions of doing so.

Acknowledgements

We would like to thank our mentor, Dr. Gokul Boyapaty, and our teacher, Mrs. Christi Sims, for their assistance in reviewing and strengthening this discussion, providing insightful comments and suggestions along the way. We would also like to thank Dr. Rick Marks for his helpful insight and provision of additional sources for our research. And last but not least, we would like to thank our families for their gracious support throughout the long process of completing this paper.

Works Cited

- American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*, 5(5). <https://doi.org/10.1176/appi.books.9780890425787>
- Blaauw, J., & Meiners, L. C. (2020). The Splenium of the Corpus Callosum: Embryology, Anatomy, Function and Imaging with Pathophysiological Hypothesis. *Neuroradiology*, 62(5), 563–585. <https://doi.org/10.1007/s00234-019-02357-z>
- Burkhardt, S. (2019, April 17). *NVLD: To Be or Not to Be Real NVLD - Children's Resource Group - A Multi-Specialty Behavioral Health Practice*. Children's Resource Group - a Multi-Specialty Behavioral Health Practice. <https://www.childrensresourcegroup.com/nvld-to-be-or-not-to-be-real-nvld/>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Fine, J. G., Musielak, K. A., & Semrud-Clikeman, M. (2013). Smaller Splenium in Children with Nonverbal Learning Disability Compared to Controls, High-Functioning Autism and ADHD. *Child Neuropsychology*, 20(6), 641–661. <https://doi.org/10.1080/09297049.2013.854763>
- Franz, C. (2000). *Diagnosis and Management of Nonverbal Learning Disorders*. <https://files.eric.ed.gov/fulltext/ED446292.pdf>
- Kouhbanani, S. S., Arabi, S. M., Zarenezhad, S., & Khosrorad, R. (2020). The Effect of Perceptual-Motor Training on Executive Functions in Children with Non-Verbal Learning Disorder. *Neuropsychiatric Disease and Treatment, Volume 16*, 1129–1137. <https://doi.org/10.2147/ndt.s252662>
- Magnus, W., Nazir, S., Anilkumar, A. C., & Shaban, K. (2023). *Attention Deficit Hyperactivity Disorder (ADHD)*. PubMed; StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK441838/>
- Maldonado, K. A., & Alsayouri, K. (2023). *Physiology, Brain*. PubMed; StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK551718/>
- Mammarella, I. C., Cardillo, R., & Semrud-Clikeman, M. (2022). Do Comorbid Symptoms Discriminate Between Autism Spectrum Disorder, ADHD and Nonverbal Learning Disability? *Research in Developmental Disabilities*, 126, 104242. <https://doi.org/10.1016/j.ridd.2022.104242>
- Mammarella, I. C., & Cornoldi, C. (2013). An Analysis of the Criteria Used to Diagnose Children with Nonverbal Learning Disability (NLD). *Child Neuropsychology*, 20(3), 255–280. <https://doi.org/10.1080/09297049.2013.796920>
- Margolis, A. E., Broitman, J., Davis, J. M., Alexander, L., Hamilton, A., Liao, Z., Banker, S., Thomas, L., Ramphal, B., Salum, G. A., Merikangas, K., Goldsmith, J., Paus, T., Keyes, K., & Milham, M. P. (2020). Estimated Prevalence of Nonverbal Learning Disability Among North American Children and Adolescents. *JAMA Network Open*, 3(4), e202551–e202551. <https://doi.org/10.1001/jamanetworkopen.2020.2551>

- Mayo Clinic. (2023, January 25). *Adult Attention-Deficit/Hyperactivity Disorder (ADHD) - Symptoms and Causes*. Mayo Clinic.
<https://www.mayoclinic.org/diseases-conditions/adult-adhd/symptoms-causes/syc-20350878>
- Miller, C. (n.d.). *How Can We Help Kids With Non-Verbal Learning Disorder?* Winston Preparatory School. <https://www.winstonprep.edu/list-detail?pk=106957>
- Otterman, D. L., Koopman-Verhoeff, M. E., White, T. J., Tiemeier, H., Bolhuis, K., & Jansen, P. W. (2019). Executive Functioning and Neurodevelopmental Disorders in Early Childhood: a Prospective Population-Based Study. *Child and Adolescent Psychiatry and Mental Health*, 13(1). <https://doi.org/10.1186/s13034-019-0299-7>
- Panos, P. T., Porter, S. S., Panos, A., Gaines, R. N., & Erdberg, P. S. (2001). An Evaluation of a Case of Agenesis of the Corpus Callosum with Rourke's Nonverbal Learning Disorder Model. *Archives of Clinical Neuropsychology*, 16(5), 507–521.
<https://doi.org/10.1093/arclin/16.5.507>
- Possin, K. L. (2010). Visual Spatial Cognition in Neurodegenerative Disease. *Neurocase*, 16(6), 466–487. <https://doi.org/10.1080/13554791003730600>
- Putnam, M. C., Steven, M. S., Doron, K. W., Riggall, A. C., & Gazzaniga, M. S. (2010). Cortical Projection Topography of the Human Splenium: Hemispheric Asymmetry and Individual Differences. *Journal of Cognitive Neuroscience*, 22(8), 1662–1669.
<https://doi.org/10.1162/jocn.2009.21290>
- Ranzenberger, L. R., & Snyder, T. (2021). *Diffusion Tensor Imaging*. PubMed; StatPearls Publishing. <https://pubmed.ncbi.nlm.nih.gov/30726046/>
- Rourke, B. P. (1987). Syndrome of Nonverbal Learning Disabilities: The Final Common Pathway of White-Matter Disease/Dysfunction? *Clinical Neuropsychologist*, 1(3), 209–234. <https://doi.org/10.1080/13854048708520056>
- Rourke, B. P., Ahmad, S. A., Collins, D. W., Hayman-Abello, B. A., Hayman-Abello, S. E., & Warriner, E. M. (2002). Child Clinical/Pediatric Neuropsychology: Some Recent Advances. *Annual Review of Psychology*, 53(1), 309–339.
<https://doi.org/10.1146/annurev.psych.53.100901.135204>
- Rylaarsdam, L., & Guemez-Gamboa, A. (2019). Genetic Causes and Modifiers of Autism Spectrum Disorder. *Frontiers in Cellular Neuroscience*, 13(13).
<https://doi.org/10.3389/fncel.2019.00385>
- Saini, S. S., Kumar, K., & Anand, A. (2023). Unveiling the Neurobiology of Specific Learning Disorders: Insights from Cognitive Neuroscience. *Annals of Neurosciences*, 30(4), 217–218. <https://doi.org/10.1177/09727531231211052>
- Salillas, E., Benavides-Varela, S., & Semenza, C. (2023). The Brain Lateralization and Development of Math Functions: Progress since Sperry, 1974. *Frontiers in Human Neuroscience*, 17. <https://doi.org/10.3389/fnhum.2023.1288154>
- The Children's Hospital of Philadelphia. (2014, August 23). *Communication Disorders*. [Www.chop.edu](http://www.chop.edu). <https://www.chop.edu/conditions-diseases/communication-disorders>

- Thompson, S. (1997). *The Source for Nonverbal Learning Disorders*. Linguisystems.
- U.S. National Library of Medicine. (2013). *Intellectual Disability: MedlinePlus Medical Encyclopedia*. Medlineplus.gov. <https://medlineplus.gov/ency/article/001523.htm>
- United States Census Bureau. (2024). *U.S. and World Population Clock*. United States Census Bureau; United States Census Bureau. <https://www.census.gov/popclock/>
- Virginia Commission on Youth. (n.d.). *Motor Disorders*.
<https://vcoy.virginia.gov/010%20Motor.pdf>
- Volden, J. (2004). Nonverbal Learning Disability. *American Journal of Speech-Language Pathology*, 13(2), 128–141. [https://doi.org/10.1044/1058-0360\(2004/014\)](https://doi.org/10.1044/1058-0360(2004/014))
- Winston Preparatory School. (n.d.). *What is the Difference Between a Non-Verbal Learning Disorder and Autism Spectrum Disorder*. Winston Preparatory School.
<https://www.winstonprep.edu/list-detail?pk=106953>

Water Pollution By Dev Shah, Jivin Mahimkar, Siddhanth Dadhich, Veer Furia, and Hriday Shah



Abstract

This research investigates the origins and health consequences of tap water pollution in suburban Mumbai, as well as suggests ways to lessen its effects. Unprocessed sewage, industrial waste, and declining water infrastructure greatly contaminate water by introducing harmful chemicals, heavy metals, and pathogens. Surveys and tests on water quality showed contamination levels surpassing allowed thresholds.

The research suggests improving infrastructure, implementing pollution regulations, and updating city water treatment facilities. Community-led efforts such as promoting affordable filters for household water and raising awareness are also emphasized. Prolonged exposure to toxins like lead and arsenic is associated with chronic diseases, including kidney damage and developmental problems in children. The results highlight the importance of working together with governments, NGOs, and citizens to guarantee long-term access to clean drinking water and protect public health.

Introduction

Background

Mumbai, one of the world's largest cities, faces severe water pollution challenges due to rapid industrialization, urbanization, and inadequate waste management. Major water bodies, including rivers and coastal areas, are heavily contaminated, threatening public health and the ecosystem.

Industrial effluents, untreated sewage, and deteriorating water supply systems are posing growing water pollution problems in the suburban regions of Mumbai. This study investigates the reasons, health effects, and potential remedies for tap water pollution in these regions.

Contaminated water harbors dangerous elements like viruses, heavy metals, and chemical residuals, resulting in widespread health problems such as waterborne illnesses and long-term disorders. This study explores the following inquiries:

1. What are the causes of water pollution in suburban Mumbai?
2. What are the impacts on its health?
3. What can be done to reduce these problems?

This research emphasizes the immediate necessity for efficient tactics to enhance water quality and safeguard public health, offering perspectives on sustainable urban water management.

Research Question

What are the levels and impacts of water pollution in different areas of Mumbai, and how can these issues be effectively mitigated?

Objectives

- To assess the current levels of water pollution in selected areas of w3 Mumbai.
- To evaluate the health and environmental impacts of water pollution.
- To identify and recommend feasible strategies for reducing water pollution.

Methodology

Data Collection Methods

- Samples (Surveys)
- Sample Size: Minimum of 50 respondents.
- Selection Criteria: Residents from diverse neighborhoods (e.g., Kandivali, Borivali, Malad) based on demographics and proximity to water bodies.
- Distribution and Collection: Surveys will be distributed via online platforms (Google Forms) and paper forms at community centers.
- Format: Surveys will be primarily written, with oral options available for those needing assistance. Translation into local languages will be provided as necessary.

Observation

- Observations: Some locations will be observed.
- Timing: Observations will occur during peak activity hours (e.g., market days, weekends).
- Participants: Research team members will conduct the observations.
- Data Capture: Observations will be documented using checklists, descriptive notes, and photographs to capture pollution sources and conditions.

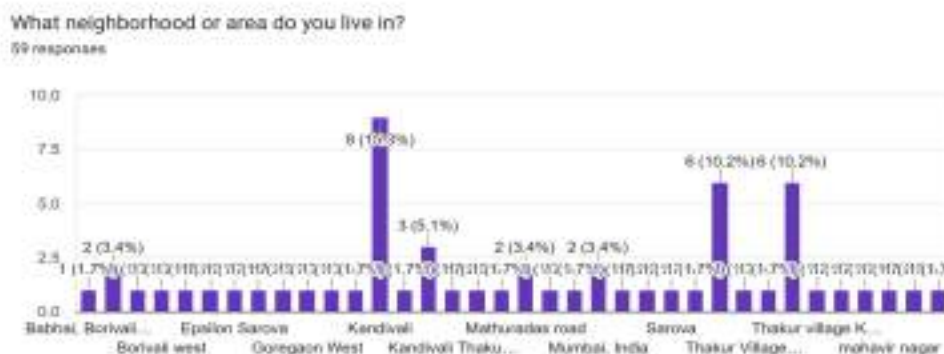
Experiments

- Repetitions: Water quality tests will be conducted at each site three times.

- **Equipment Needed:** Water quality testing kits (for parameters such as pH, turbidity, and heavy metals).
- **Data Capture:** Data will be recorded in a digital format.
- **Data Collection Tools**
- **Interview Questions**
 1. What do you perceive as the main sources of water pollution in your area?
 2. How has water pollution affected your health and the community?
 3. What measures do you believe should be implemented to reduce pollution?
- **Survey Questions**
 1. How concerned are you about water pollution in your area? (1-5 scale)
 2. Have you experienced any health issues attributed to water pollution? (Yes/No)
 3. What actions would you support to improve water quality?
- **Observation Collection Sheet:** A standardized form documenting:
 1. Location
 2. Date and time
 3. Observed pollution sources
 4. Visual documentation (photos)
- **Special Equipment:** Water quality testing kits, noise level meters, mobile apps for data collection.
- **Experimental Procedure:** Detailed protocol for sampling and testing water quality, including safety measures and calibration of equipment.
- **Secondary Data Sources:** Local government reports, NGO publications, environmental studies.

Findings

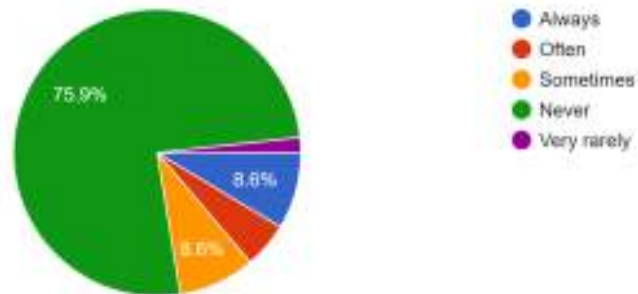
Surveys



Most of the people are from Kandivali, Borivali and Malad .

How often do you use tap water directly for drinking?

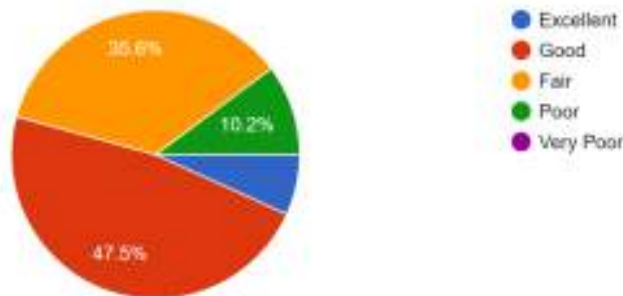
58 responses



As the pie chart suggests, around 75% of the respondents do not use tap water directly.

How would you rate the quality of tap water in your area?

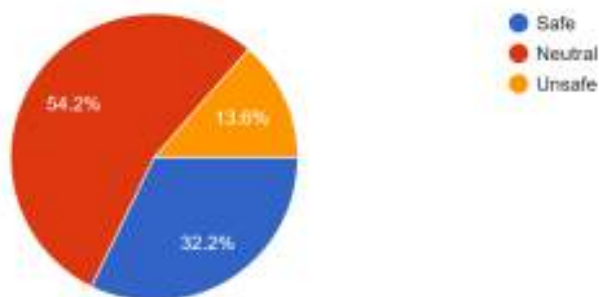
59 responses



Most of the respondents feel the water quality is between good and fair, which is not a very good sign.

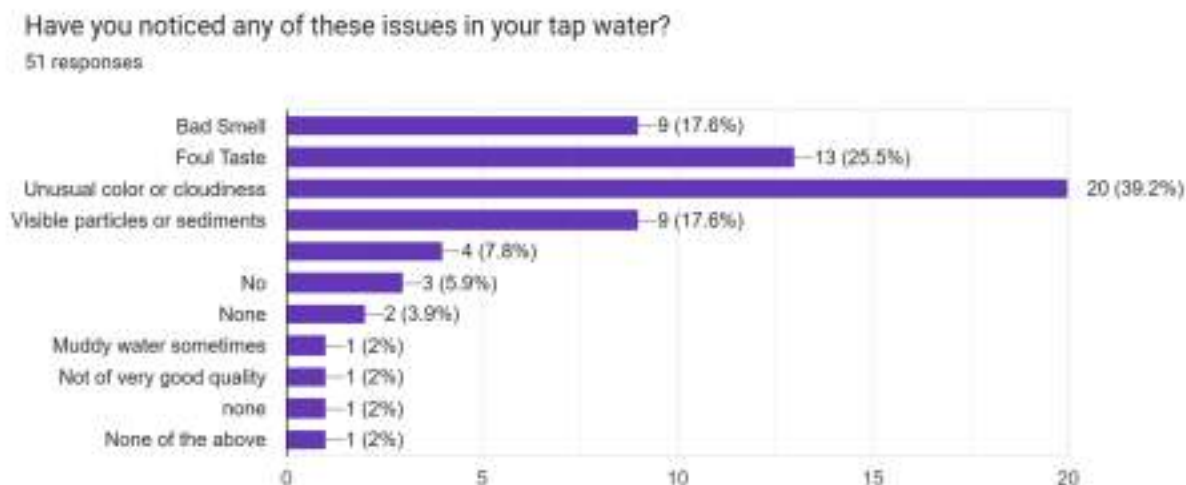
How safe do you feel using tap water for daily activities (e.g., drinking, cooking, bathing)?

59 responses



Most of the respondents feel neither safe nor unsafe while using the tap water, but this can be

improved by taking extra precautions regarding this.



The majority of the respondents faced many issues with their tap water, foul taste, unusual colour or cloudiness etc. which is not a great sign for a big and developed city like Mumbai.

Additional Comments: To address the sources of pollution, community comment on water quality emphasizes the necessity of group efforts, with a focus on waste segregation, cleanup drives, and environmentally acceptable methods. Some responders support routine pipeline maintenance, higher water quality standards, and better waste disposal systems. To guarantee long-lasting benefits, concerns over the smell of recycled water were raised, along with recommendations for more aggressive government initiatives and awareness-raising activities. The remarks often show a mixture of happiness and a strong need for improved community involvement and water management.

Observations

Field surveys at major water sources (such as Vihar Lake, Tulsi Lake, and Bhatsa Lake) revealed particular sources of pollution.

Industrial Discharge: Factories near Bhatsa Lake were observed releasing untreated wastewater into nearby streams that eventually flow into the lake.

Plastic Debris: A significant buildup of plastic and non-biodegradable waste was found around Vihar Lake due to littering and illegal dumping from nearby settlements.

Sewage Overflow: Encroachments and poorly maintained drainage systems near Vihar and Tulsi Lakes resulted in frequent sewage spills into the lakes, especially during the monsoon season.

Agricultural Runoff: Farms around Tansa and Modak Sagar Lakes contributed fertilizers and pesticides to the water bodies through runoff, impacting water quality.

Encroachments: Illegal settlements near Vihar and Tulsi Lakes were noted to discharge domestic

wastewater and solid waste directly into the lakes.

Tourist Activities: Human activity in Sanjay Gandhi National Park contributed to localized pollution around Tulsi Lake, including littering and waste from visitors.

General Findings: While remote lakes like Bhatsa and Vaitarna face lower pollution, lakes closer to urban and suburban areas—such as Vihar and Tulsi—are significantly impacted by sewage, plastic, and human encroachments.

Environmental Effects Observed

Observations showed obvious evidence of ecological harm, such as:

- In contaminated waterways, fish die.
- Algal blooms are a sign of eutrophication or nutrient contamination.
- Water discolouration and unpleasant smells discourage recreational use.

Testing for Water Quality

Important Results

The following results were obtained from tests that were carried out at three distinct times and locations. In Kandivali, we tested the water of Thakur village. And some more areas of Borivali, Malad and Goregaon.

- pH levels varied from 6.0 to 9.0, occasionally exceeding permissible limits.
- Turbidity: Water clarity was decreased in locations close to industrial effluents by elevated turbidity levels (over 5 NTU).
- Heavy Metals: In several samples, detectable amounts of chromium (0.02 mg/L) and lead (up to 0.04 mg/L) were higher than the allowable limits established by BIS and WHO guidelines.

Times of Peak Pollution

Post-monsoon pollution levels were consistently higher, most likely as a result of untreated garbage entering aquatic bodies through urban runoff.

Interviews

Interviewees highlighted persistent problems like:

- Lack of responsibility for the disposal of industrial waste.
- Waste treatment plants in the municipality are overworked.
- Inadequate public knowledge of appropriate trash segregation procedures.

Effects on Health

Respondents mentioned common ailments that are directly related to contaminated water sources, including skin conditions and gastrointestinal disorders.

Possible Solutions

❖ *Upgrade Water Infrastructure:*

Replace aging pipes with modern materials to prevent leaks and contamination. For example, *Mumbai's Bhandup Water Treatment Plant*, Asia's largest, could benefit from further upgrades to improve efficiency and filtration processes. This would minimize contamination risks during distribution.

- *Example:* Cities like *Singapore* have modernized their entire water infrastructure to reduce waste and contamination through advanced recycling systems.

❖ *Strict Industrial Regulation:*

Enforce penalties and incentives for industries to adopt eco-friendly waste management. For instance, the *Maharashtra Pollution Control Board (MPCB)* already monitors industries, but more rigorous implementation is needed to prevent effluent discharge into lakes like *Vihar* and *Bhatsa*.

- *Example:* *CETPs (Common Effluent Treatment Plants)* used in Gujarat's industrial zones have shown effectiveness in controlling industrial water pollution.

❖ *Community-Led Awareness Programs:*

Engage local communities through workshops and campaigns to promote waste reduction and safe disposal practices. For instance, the *Clean Powai Lake Initiative* in Mumbai involved residents in cleanup drives, significantly improving the lake's condition.

- *Example:* The *Swachh Bharat Mission* has shown success in waste segregation awareness at the grassroots level.

❖ *Affordable Household Water Filtration:*

Distribute cost-effective solutions such as ceramic or UV water purifiers in areas lacking access to treated water. NGOs like *WaterAid India* have implemented similar programs in rural areas to ensure clean drinking water.

- *Example:* *LifeStraw*, a portable water filter, has been successfully used in resource-limited settings globally to prevent waterborne diseases.

❖ *Real-Time Water Quality Monitoring:*

Use IoT sensors to monitor and track water pollution. For instance, installing sensors in lakes like *Vihar* and *Tulsi* could help detect sewage inflows or chemical contamination promptly.

- *Example:* Cities like *Bangalore* have piloted IoT-based water monitoring in reservoirs to respond to contamination early.

❖ *Eco-Restoration Initiatives:*

Promote natural filtration through wetland creation and vegetation buffers around lakes. For example, introducing *phytoremediation plants* in catchment areas could reduce nutrient runoff and control algal blooms in lakes like *Powai* and *Tansa*.

- *Example:* The *Yamuna Biodiversity Park* in Delhi demonstrates how degraded ecosystems can be restored with community and governmental efforts.

Conclusion

According to this study, the main causes of water pollution in suburban Mumbai are untreated sewage, industrial waste, and antiquated infrastructure. This poses a major concern to public health, since it may result in heavy metal poisoning and other waterborne ailments. Alarming pollution levels, ecological harm, and the urgent need for a multipronged strategy to address these issues are highlighted in the report. Enforcing stronger industrial waste rules, repairing outdated pipes, and modernizing municipal water treatment systems are among solutions. Participation from the community is similarly important, with a focus on waste management education, reasonably priced home filtration devices, and cooperation between residents, NGOs, and local government. These actions are necessary to prevent environmental damage, safeguard public health, and guarantee long-term access to clean water.

Works Cited

<https://timesofindia.indiatimes.com/>

<https://www.coolgeography.co.uk>

<https://timesofindia.indiatimes.com/>

<https://bmcpublichealth.biomedcentral.com/>

<https://www.weforum.org/>

Ecclesiastical and Classical Latin: A Comprehensive Study of Linguistic Evolution and Theological Implications By Suya Fang

Abstract

The Latin language underwent significant changes from its Classical form to its Ecclesiastical form. This paper explores the key differences between Ecclesiastical Latin and Classical Latin, focusing on pronunciation, vocabulary, and grammar. It examines how these changes reflect the broader historical transition from the Roman Empire to the Christian Middle Ages. The study also highlights the theological implications of these linguistic changes, exploring how the Catholic Church shaped Ecclesiastical Latin as a tool for religious and cultural unity in medieval Europe. Despite its historical significance, Latin is known as a dead language, making it difficult to find reliable online sources that explain its evolution. Many existing resources either oversimplify the distinctions between Classical and Ecclesiastical Latin or fail to address their deeper cultural and theological implications. This paper seeks to bridge this gap by providing a detailed analysis. Latin, though often seen as a language of the past, remains an essential key to understanding the foundations of Western civilization. As a study in historical linguistics, this paper is based on recurring information from research articles and books. While linguistic interpretations may differ, this analysis presented reflects widely accepted scholarly perspectives—such as those discussed in *Late Latin and Early Romance in Spain and Carolingian France*.

Introduction

Classical Latin, the language of the Roman elite, was once an influential tool in politics, law, and literature. However, as the Roman Empire declined and the Catholic Church rose during the Middle Ages, this language evolved from Classical Latin (CL) to Ecclesiastical Latin (EL). This linguistic shift mirrors the broader historical and cultural transitions of the period. Classical Latin, spoken during the Roman Republic and Empire, was the language of intellectuals, orators, and philosophers, while Ecclesiastical Latin emerged as the language of the Christian Church. The transition from Classical to Ecclesiastical Latin was not merely a linguistic evolution but also a reflection of the social, political, and theological imperatives of the time. Ecclesiastical Latin, although rooted in Classical Latin, was significantly shaped by the Church and supports its evolving liturgical practices and expanding Christian community (Mohrmann, 1959).

To understand this linguistic evolution, it is essential to examine the historical background in which Latin was transformed. Early Christian communities in Rome, predominantly Greek-speaking, utilized Greek in early Christian texts and inscriptions (Hurtado, 2008). However, as Christianity spread throughout the Roman Empire, Latin gradually replaced Greek as the language of Christian worship in the Western Church. This shift reached its peak with St. Jerome's translation of the Bible into Latin in the late 4th and early 5th century. Rather than creating an entirely new Latin Bible, Jerome revised existing Old Latin translations, particularly for the Gospels and Psalms, while incorporating Greek and Hebrew influences.

Jerome's translation played a crucial role in shaping the Latin used in the Church for centuries (Bruce, 1988).

The Church's adoption of Latin serves both practical and theological purposes. Latin was the language of the Western Roman emperor; therefore, it carried authority and universality in the West. Latin's authority and universality allowed the Church to continue its mission of spreading Christianity across Europe. Additionally, Latin helped the Church maintain doctrinal consistency and counter regional variations in religious practice (Lafferty, 2003).

In this paper, we will explore how the phonological, lexical, and grammatical changes in Ecclesiastical Latin reflect the broader theological and cultural transformations of the medieval period. By examining these linguistic changes, we will discover how Latin's evolution helped shape the religious and social framework of medieval Europe.

I. Pronunciation Differences Between Classical and Ecclesiastical Latin

The pronunciation of Classical Latin was different from the phonetic structures that emerged in the Romance languages. Linguists emphasize that in Classical Latin, vowels had distinct long and short forms, and diphthongs such as 'ae' were pronounced as separate sounds, as in *caelum* (/ˈkaɛ.lum/). Consonants were also distinct: for instance, 'c' and 'g' were always hard, regardless of the following vowel. Additionally, the letter 'v' was pronounced as a /w/ sound, as in 'vini' (wine), which was spoken /wi.ni/ (Allen, 2004). The clarity and precision of Classical Latin pronunciation were essential in Roman rhetoric and public discourse (Palmer, 1999). Cicero's speeches, for example, relied on the distinct pronunciation of vowels to show subtle emphases and meanings.

In his *Pro Archia Poeta*, Cicero used the distinct vowel lengths in Classical Latin to add emphasis and subtlety to his argumentation. Consider the opening line:

"Si quid est in me ingenii, iudices, quod sentio quam sit exiguum..."
("If there is any talent in me, judges, which I feel is very slight...").

Here, the long vowel in 'ingenii' (/inˈge.ni.i/) creates a drawn-out sound that emphasizes the idea of "talent." This elongation draws attention to Cicero's humility and skill, while the long 'u' in 'exiguum' (/ekˈsi.gu.um/) adds weight to the adjective "slight," reinforcing his modest self-presentation. The placement of 'exiguum' at the end of the clause creates a natural pause, allowing the audience to reflect on Cicero's humility.

Another example of the importance of Classical Latin pronunciation in writing can be found in Virgil's *Aeneid* (Virgil: *Aeneid* I., 2020).

In the opening line: "Arma virumque canō, Trōiae quī prīmus ab ōrīs"
("I sing of arms and the man, who first from the shores of Troy")

The long vowels in 'virumque' and 'prīmus' create a rhythmic cadence that enhances the epic tone of the poem. The careful distinction between long and short vowels allowed poets like Virgil to craft verses with precise metrical patterns—a hallmark of Classical Latin literature.

Ecclesiastical Latin, however, focused less on the rhetorical effects but rather on practicality. It underwent phonetic changes. The letter ‘c’ before ‘e’ and ‘i’ was softened to a /tʃ/ sound, as in ‘civitas’ (city), pronounced /ˈtʃi.vi.tas/ in Italian Ecclesiastical Latin, compared to the hard ‘c’ in Classical Latin (Herman, 2000). Likewise, the letter ‘v’ began to be pronounced as it is in modern Italian, which is closer to the /v/ sound used in English rather than the /w/ sound in Classical Latin. Vowel sounds were also simplified: diphthongs like ‘ae’ and ‘oe’ became monophthongs (Allen, 2004).

However, these shifts were not uniform across all regions. Germanic areas, for instance, retained hard ‘c’ pronunciations longer. This showed the uneven adoption of Ecclesiastical Latin’s phonetic rules in different areas. These changes were influenced by the regional dialects of Vulgar Latin—the spoken Latin of the common people—that diverged in the western provinces of the Roman Empire (Palmer, 1999). With the creation of Ecclesiastical Latin, Latin shifted from a primarily spoken language, due to the collapse of the Roman Empire, to a liturgical and scholarly tool in the Christian Church, allowing regional features to shape Ecclesiastical Latin. This caused the change in pronunciation and reflected the Church’s desire to make Latin more accessible to a broader audience. By adopting phonetic patterns that were closer to the vernacular languages of medieval Europe, the Church ensured that Latin could be understood and used by both clergy and common folk.

II. Vocabulary Shifts

The primary spread of Classical Latin was driven by Rome’s legal system and military, emphasizing concepts of power, governance, and law (Darnell, 2024). As a result, Classical Latin reflected Rome’s highly organized and militarized society, where terms like ‘legio’ (legion) and ‘ius’ (law) were central to both Roman legal texts and military operations. These terms appear in the works of intellectuals such as Cicero and Virgil, who used them to explore philosophical and political ideas. For example, in Caesar’s *Commentarii de Bello Gallico*, the term ‘legio’ is used repeatedly to describe the Roman military units that conquered Gaul. Similarly, ‘ius’ appears frequently in Roman legal texts, such as the *Twelve Tables*, revealing the importance of law in Roman society.

Ecclesiastical Latin, in contrast, underwent a significant transformation in vocabulary to meet the theological and liturgical needs of the Church. New terms were either borrowed from other languages or repurposed to express Christian doctrine. A notable trend was the Latinization of Greek-origin words, particularly those used in sacred texts (Connaughton, 2008). Terms, such as ‘sacramentum’—originally meaning “oath” in Latin—were adapted to express Christian doctrine. This symbolically aligned Christians with soldiers swearing allegiance—not to the emperor, but to Christ. By the fourth century, Augustine’s sermons popularized sacramentum to frame baptism as a binding covenant, thus recontextualizing the term from its martial origins and sacralizing it within ecclesiastical authority. These vocabulary shifts not only reflect linguistic evolution but also highlight the profound cultural and religious transformation from a Roman imperial worldview to a Christian theological framework. The adaptation of words ensured that

Christian doctrine could be spread within the Latin-speaking world while retaining continuity with classical traditions (Brakke & Deliyannis, 2016).

III. Grammatical Evolution

Classical Latin relied on a complex system of noun declensions and verb conjugations to convey meaning. The precision of Classical Latin grammar allowed for a high degree of syntactic flexibility, enabling authors to construct sentences with nuanced meanings and intricate structures. Ecclesiastical Latin, however, saw a simplification in some areas, particularly in syntax and word order. This simplification was influenced by the spoken forms of Latin, often referred to as Vulgar Latin, which gradually diverged from the highly stylized literary Latin. Verb usage also changed. The Classical Latin subjunctive, widely used for hypothetical or subordinate clauses, had a more variable use in Ecclesiastical Latin (Wheelock, 2011). More direct and simplified sentence structures became prevalent, aligning with the Church's goal of making religious texts more accessible to a broader audience (Herman, 2000). Another key development was the increased use of the future passive periphrastic construction to express necessity or obligation, a shift that reflected the theological tone of many Ecclesiastical Latin writings. Additionally, word order became more rigid, moving closer to the Subject-Verb-Object structure found in many modern Romance languages, as opposed to the relatively flexible word order in Classical Latin.

IV. Theological Implications of Linguistic Changes

The evolution of Latin from its Classical to Ecclesiastical form was not merely a linguistic phenomenon but also had profound theological implications. The shifts in Ecclesiastical Latin reflected a movement toward clarity, universality, and accessibility—qualities essential for a religion seeking to reach all people, regardless of social class or education level. The transition to Ecclesiastical Latin also reinforced the authority of the Catholic Church. By standardizing a specific form of Latin for liturgical and doctrinal use, the Church ensured linguistic unity that facilitated the spread of religious teachings (Mohrmann, 1959). This linguistic unity helped maintain consistency in theological interpretations, preventing regional languages from introducing doctrinal variations, a challenge that Catholic teachings in Greek faced (Connaughton, 2008). This strategy proved effective, as it led to the Vulgate becoming the definitive version of the Bible for centuries, shaping Christian theology.

The Church's linguistic standardization was a form of doctrinal warfare. When the Council of Tours (813 CE) mandated vernacular sermons but retained Latin Masses, it elevated Latin as a sacred language—a defense against regional heterodoxy. The duality of Latin—its accessibility and connotated power—shaped medieval Europe's religious identity. Moreover, the use of Ecclesiastical Latin in Church rituals reinforced the perception of Latin as a sacred language. The association of Latin with the divine contributed to its elevated status in religious settings. This perception helped Latin become a unifying language within Western Christendom, bridging geographical and cultural divides (Lafferty, 2003).

Conclusion

The evolution from Classical Latin to Ecclesiastical Latin is not only a linguistic shift but also a reflection of broader historical, cultural, and religious changes. As the Roman Empire weakened and Christianity spread across Europe, Latin adapted to meet the needs of a changing world. The phonological, lexical, and grammatical changes in Ecclesiastical Latin mirror the influence of the Catholic Church as it became the central institution in medieval Europe, uniting diverse communities under a common religious and linguistic framework. The linguistic evolution of Latin also has profound theological implications. By adopting and modifying the language to express Christian concepts and facilitate worship, the Church not only helped unify Europe through a shared liturgical language but also reinforced its role as the mediator between the divine and the human. Through these linguistic transformations, Latin became not only the language of intellectual discourse but also of faith, reflecting the shift in European society from paganism to Christianity.

The transition from Classical Latin to Ecclesiastical Latin reveals more than just the evolution of a language; it discusses the broader historical and theological forces that shaped medieval Europe. Understanding this linguistic transformation provides valuable insights into the intersection of language, culture, and religion during one of the most important periods in Western history.

Works Cited

- Auerbach, E., & Trask, W. R. (2013). *Mimesis : the representation of reality in western literature*. Princeton University Press.
- Brakke, D., & Deliyannis, D. (2016). *Shifting Cultural Frontiers in Late Antiquity*. Routledge.
- Bruce, F. F. (1988). *The Canon of Scripture*. Inter-Varsity Press.
- CHURCH FATHERS: *De Corona (Tertullian)*. (n.d.). Wwww.newadvent.org. Retrieved March 25, 2025, from <https://www.newadvent.org/fathers/0304.htm>
- Connaughton , M. (2008). *Latin Allocution and the Applications and Usage of Latin as a Modern Language by the Vatican City State* .
https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send?accession=muhonors1111069773&disposition=inline
- Darnell, Lilianna, (2024). "The Roman Language Policy: Its Parts, Presence, and Consequences." Honors Bachelor of Arts. 54. <https://www.exhibit.xavier.edu/hab/54>
- Hurtado, L. W. (2008). *The Earliest Christian Artifacts : Manuscripts and Christian Origins*. Eerdmans.
- Lafferty, M. (2003). Translating Faith from Greek to Latin: Romanitas and Christianitas in Late Fourth-Century Rome and Milan. *Journal of Early Christian Studies*, 11(1).
http://www.shirleymohr.com/JHU/Sample_Articles_JHUP/ERL_2003_11_1.pdf
- Mohrmann, C. (1959). *Liturgical Latin: Its Origins and Character*.
- Palmer, L. R. (1999). *The Latin language*. Bristol: Bristol Classical Press.
- The Speech on Behalf of Archias the Poet: Chapter I*. (2025). Loeb Classical Library.
https://www.loebclassics.com/view/marcus_tullius_cicero-pro_archia/1923/pb_LCL158.7.xml
- Vergil: Aeneid I*. (2020). Thelatinlibrary.com. <https://www.thelatinlibrary.com/vergil/aen1.shtml>
- Wheelock, F. M., & Lafleur, R. A. (2011). *Wheelock's Latin : [the classic introductory Latin course, based on the writings of Cicero, Vergil, and other major Roman authors]*. Collins Reference.
- Allen, William (2004). *Vox Latina: a guide to the pronunciation of classical Latin*. Cambridge Univ. Press.

Emerging Middle Powers: The Geopolitical Strategies of Türkiye and Azerbaijan Amidst the Russo-Ukrainian War

By Emil Javanshir

Abstract

Over the past few decades, a new category of states has emerged within the global order, known as middle powers. These nations hold significant geopolitical influence, often measured by factors like GDP and military strength, while also distinguishing themselves through a strong commitment to multilateralism in their international engagements. While in the field of international relations the theory behind middle powers is fairly contemporary, the rapidly growing geopolitical influence that these nations yield warrants an investigation into the common practices of middle powers in an increasingly polarised twenty-first century. Although the geopolitical strategies of middle powers vary across multiple variables, the foreign policy of these like-minded states exhibits a noticeable pattern during periods of exogenous shocks, such as the Russian invasion of Ukraine commencing in February of 2022, which sent sizable political shockwaves throughout the geopolitical landscape. This paper focuses on the reactions of two middle powers, Türkiye and Azerbaijan through the review of academic sources as well as the analysis of original data, including a written interview conducted with a Senior Azerbaijani Diplomat representing the Ministry of Foreign Affairs of Azerbaijan. These findings demonstrate that middle powers favour the practice of hedging political risks between two warring blocs when not threatened by the consequences of escalation.

Introduction

The era of bipolarity, where global geopolitics was dominated by two rival superpowers, has gradually faded into history. Following the collapse of the Soviet Union in 1991 and the subsequent unipolar moment led by the United States, middle powers have increasingly expanded their influence, shaping world politics through strategic diplomacy, economic growth, and multilateral engagement. While international relations scholars have proposed different avenues for classifying a state as a middle power, they are usually referred to as states that, while not as dominant as great powers (i.e., UN Security Council permanent members) still exert a considerable influence over global politics. Over the past few decades, emerging middle powers have become more assertive and outspoken in their actions on the world stage, whether it be through multilateral agreements organised by international organisations, or through unilateral endeavours. While the weight of middle powers has clearly risen in recent years, existing literature on this subset of states is still underdeveloped and does not answer critical questions regarding the behaviour of middle powers vis-à-vis change and insecurity in the international system. In a world of frequent and severe exogenous shocks, such as the Russian invasion of Ukraine, how do middle powers react? As middle powers emerge as an increasingly influential cohort of the world order, it is more imperative than ever to conduct comprehensive research on how rising global tensions and intensified armed conflict shape the foreign policy of middle

powers and how these changes in geopolitical strategy affect regional stability, global power dynamics, and the evolving paradigms of international relations.

As the USA pivots its geopolitical focus away from Europe toward East Asia and the emerging superpower China, many traditional middle powers, particularly those in Eastern Europe, have shouldered the responsibility of providing Ukraine with vital military and humanitarian aid. Indeed, the efforts of European states have been immensely influential in aiding Ukraine to ward off Russian imperialism. This highlights a key difference between middle powers: those which are associated with the traditional Western vs Eastern rivalry, and nations which participate in a precarious balancing act between these two blocs, nations which tend to be ‘emerging’ middle powers. However, this leads to an intensification of regional polarity as countries align themselves with a faction. This heightens the likelihood of an escalation into a widespread regional conflict, or in the worst-case scenario, expanding into global confrontation, thus dragging the global heavyweights America and China into total warfare. While middle powers in the Western camp have adamantly denounced the Russian invasion, which violates international law, and cemented their support for Ukraine, others, particularly those linked to Russia through trade deals, political unity or historical ties, appear more reluctant to align themselves with one of the warring sides.

This research paper will delve into two such cases. Firstly, the Republic of Türkiye, which has historically been a bitter military rival to Russia since 1568 until the normalisation of relations following World War One. Therefore, this paper investigates the cause of Türkiye’s hesitance to align with sanctions against Russia. Türkiye’s foreign policy very accurately represents the unique nature of the nation’s political standing. Once a dominant empire in Southeastern Europe, the Ottomans’ geopolitical influence was significantly diminished come the close of World War I. Following the founding of the Republic of Türkiye in 1923, the nation has progressively shifted to a more Western-orientated structure in terms of industry, economy and society. However, Türkiye’s foreign policy remains fairly ambiguous and perhaps even inflammatory and confrontational, especially in the Eastern Mediterranean. This places Türkiye in a delicate position, balancing between Russia and the bloc of European nations supporting Ukraine. This relationship will be further explored as well as how Türkiye balances their affiliation with either side.

The second case of a middle power trapped between conflicting interests in Ukraine is Azerbaijan. Azerbaijan shares many cultural ties with Russia with the majority of Azerbaijanis being proficient in Russian due to the Soviet Occupation of Azerbaijan between 1920-1991. Since the fall of the USSR, Russo-Azerbaijani relations have remained relatively friendly. However, in an interview with a senior Azerbaijani diplomat representing the Ministry of Foreign Affairs, it is apparent that the Azerbaijani government are adamant in their support for Ukraine as ‘from the very first day of the ongoing conflict in Ukraine, Azerbaijan has consistently expressed its support for the sovereignty and territorial integrity of Ukraine within its internationally recognized borders (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025)’. Moreover, Azerbaijan has been

inching closer and closer to a stable and positive relationship with the EU and Europe as a whole, whether it be through signing new gas deals (EU steps up renewable energy cooperation with Azerbaijan”, European Commission, 2025) or hosting COP 29.

Interestingly, both Azerbaijan and Türkiye are regarded as maintaining one of the best relations in the world due to deep historical and cultural ties. However, this does not necessarily mean their geopolitical strategies in response to the Russian invasion of Ukraine will be similar.

Identifying Middle Powers: The case of Türkiye and Azerbaijan

Existing literature on middle powers will be useful in providing a theoretical framework for the remainder of the paper. This will be achieved by delving into two sources: *Middle Powers* by Marion Laurence and *One Hundred Years of Turkish Foreign Policy in Eastern Mediterranean(1923–2023): A Neoclassical Realist Analysis of Turkey as a Middle Power* by Bezen Balamir Coskun. A comprehensive review of this literature will provide a foundation for middle power theory by initially reviewing the definition of a middle power before analysing the general trend observed in the geopolitical strategies of middle powers. Subsequently, there will be analysis of Türkiye’s historical foreign policies which will be employed to formulate a hypothesis on the reactions of Türkiye and Azerbaijan to Russia’s invasion of Ukraine.

What defines a middle power? The etymology of the term dates to the Congress of Vienna though was first implemented by Canada who ‘used it to distinguish themselves from smaller states and secure a relatively favorable position in the postwar order’ (Laurence, 2023). This dynamic of self-representation by middle powers will be further explored. Laurence also notes that classification of middle powers is ‘inherently relational’ (Laurence, 2023) meaning that despite any judgement criteria of what constitutes a middle power, they must lack one or a number of characteristics in comparison to ‘great powers’ while being superior to ‘small powers’ when comparing other hallmarks of a country. The definitive definition of a middle power is still widely contested by scholars, however, there are three central methods of categorising middle powers which have come to the forefront of debate, ‘one positional, one behavioural and one identity-based’ (Carr, 2014). Before embellishing on the approaches to defining middle powers, it is important to note the overarching connotations of the term. The label ‘middle powers’ is inherently perceived as being associated with positive underlying connotations whether it be through insinuating that a nation favours multilateralism or is dedicated to neutrality and diplomacy.

The first method to categorize middle powers outlined by Laurence is the ‘positional approach’. An approach which entails ranking nations through the use of quantitative data in order to establish a hierarchy of nation states. Laurence notes how ‘the concepts of strength and hierarchy are central to this definition’ (Laurence, 2023). When partaking in a comprehensive positional approach, several key indicators such as GDP, total population, military rankings or landmass will be measured and summed together to provide a total score. This will facilitate comparative analysis of each country’s position in the global order which can then further be dissected into separate sections classifying whether a country is a small, middle or great power.

However, there are several challenges to performing a positional approach when classifying middle powers. Firstly and most significantly, it is complex to determine the appropriate weight to assign to each indicator. As Laurence questions: “Does military spending matter more than GDP per capita? Is having a medium-sized population more important than geographic size?” (Laurence, 2023). Some scholars have attempted to offset this problem by encompassing a wider range of more niche identifiers ‘like geographic location, technological and diplomatic capability, and qualitative indicators like national reputation’ (Laurence, 2023). Another flaw of the positional approach is that even if an objective ranking is achieved and a hierarchy is established, the question of how the rankings should be divided when allocating these geopolitical labels remains unresolved. What is the threshold between small and middle powers and middle and great powers?

The behavioural approach attempts to resolve this limitation posed by measuring quantitative data by instead observing ‘how middle powers actually comport themselves in international politics’ (Laurence, 2023). This approach entails identifying patterns in the behaviours of middle powers on the global stage and observing the political attitudes of nations to determine if they resemble these values. This raises the question: Which specific behaviors can be considered definitive of middle powers' foreign policy conduct? Advocates of the behavioural approach, Cooper, Higgott and Nosal argue that middle powers have the unique ‘tendency to pursue multilateral solutions to international problems, their tendency to embrace compromise positions in international disputes and their tendency to embrace notions of ‘good international citizenship’ to guide their diplomacy (Cooper et.al., 1993).” This hypothesis is based on studies of Canada and Australia which are typically viewed as the first widely acknowledged middle powers. Scholars argue that Canada and Australia resemble a typical middle power because their “managerial capacities —especially their skills in negotiation and “mediatory diplomacy”—are among their greatest contributions to international politics (Cooper, 1997).” Behavioural definitions have been criticized because they emulate the view of a very western centric norm of the behaviour of middle powers. This is followed up by Laurence as they explain that “countries like South Korea or Indonesia—do not actually follow the foreign policy behaviors theorized by the mainstream literature on middle powers (Laurence, 2023).” The increasing prevalence of ‘emerging’ middle powers spans several continents and cultures. Therefore, the obsolete standard of the multilateral tendencies which middle powers exhibit limits the effectiveness of the behavioural approach.

The identity-based approach is one which refrains from incorporating quantitative or qualitative data. This definition categorises middle powers on the basis of self-identification as a middle power. Initially employed by Canada following the second World War in order to ‘lobby for greater influence within postwar institutions like the United Nations’, the identity based approach is the most rudimentary method of defining middle powers. However, critics argue that self-identification of a middle power tends to have ulterior motives to creating a hierarchy in the global order as the term is a ‘deliberately constructed “political category”’ (Carr, 2014), one that

is by design laden with positive connotations.’ Therefore, it is imperative to deduce if a country is self-identifying as a middle power as a form of political marketing.

It is evident that classification of middle powers is a very nuanced topic and that opinions on the true definition of a middle power may vary drastically according to several different scholars. This paper will incorporate a hybrid approach to defining middle powers, one that combines the quantitative material measures of the positional approach, the analysis of a nation’s conduct on the global stage, and to a lesser extent, self-identification practices of nations claiming to be a middle power.

Applying this criterion to the cases of Türkiye and Azerbaijan, it will become apparent that while middle powers do broadly possess similar attributes, there is a varying spectrum of middle powers particularly when constructing a hierarchy based upon power and international influence. Türkiye appears to constitute a textbook middle power when taking into account quantitative data. It ranks highly when looking at raw economic output, being the 17th and 12th largest economy globally by nominal GDP (Nominal gross domestic product (GDP), OECD, 2025) and GDP PPP (GDP, PPP (current international \$)’’, World Bank Group, 2025) respectively. It boasts the second largest population in Europe and the 9th strongest military in the world (GDP, PPP (current international \$)’’, World Bank Group, 2025). Türkiye is unequivocally a regional powerhouse and its presence stretches from the Northernmost point of Europe down to the Middle East. However, Türkiye ranks modestly in other indicators such as GDP per Capita, HDI, unemployment and innovation. Assessing Türkiye’s conduct on the global stage does illustrate some tendencies associated with middle powers. Türkiye is a key member of NATO and the UN as well as being a key strategic partner to the EU. However closer to home, Türkiye’s foreign policy often entails political conduct which is frequently labelled as aggressive and inflammatory particularly in the Eastern Mediterranean and Aegean Sea and in Syria. For example, prior to the collapse of the Assad government, Türkiye had frequently clashed with the Syrian Democratic Forces (SDF), the Syrian army and to an extent with the Russian Air Force, due to the infamous downing of a Russian Sukhoi SU-24 fighter bomber by Turkish F16s near the Turkish-Syrian border. This highlights Türkiye’s frequent enactment of unilateral policy, particularly when it comes to military actions. However, on balance, Türkiye’s actions on the wider geopolitical stage generally conform to the norm exhibited by other widely established middle powers such as Canada and Australia.

Azerbaijan is not dissimilar to Türkiye in its role as a regional power in the Caucasus although its wider geopolitical influence is undoubtedly significantly less extensive than Türkiye’s. When considering quantitative indicators, Azerbaijan is observed to be towards the middle to lower ranking of the hierarchy. In terms of GDP, it ranks 83rd (Nominal gross domestic product’’ (GDP). OECD, 2025) and 71st (GDP, PPP (current international \$)’’, World Bank Group, 2025) when measuring nominal GDP and GDP PPP respectively. However, Azerbaijan possesses significant political and economic leverage thanks to its large reserves of crude oil and natural gas. In terms of political conduct, Azerbaijan is becoming increasingly active in regional multilateral organisations such as GUAM, an economic forum including the four ex-Soviet

countries Georgia, Ukraine, Azerbaijan and Moldova. Even in an economic landscape, Azerbaijan engages in multilateralism through its partaking in OPEC Plus, a global oil cartel. After conducting this analysis, it is evident that Türkiye and Azerbaijan are middle powers albeit on different regions of the spectrum and will therefore also be referred to as ‘emerging middle powers’.

The foreign policy of Türkiye and whether it represents the typical foreign policy of a middle power is a very nuanced and complex argument. To provide a better understanding of Türkiye’s historical geopolitical strategies in order to develop a hypothesis for Türkiye’s reaction to the Russo-Ukrainian war, *One Hundred Years of Turkish Foreign Policy in Eastern Mediterranean (1923–2023): A Neoclassical Realist Analysis of Turkey as a Middle Power* by Bezen Balamir Coskun will be briefly reviewed. The main focal point of Türkiye’s foreign policy has been the Eastern Mediterranean where Türkiye employs its Blue Homeland Doctrine. During the interwar period, ‘Turkey played on the balance of power between Britain-France, Germany-Italy, and the USSR’ (Coskun, 2023) a concept known as hedging. This method of geopolitical conduct applies when a state maintains stable and positive political ties with two separate blocs of conflicting parties while acting as an intermediary between them, a key characteristic of middle powers. Interestingly, Türkiye has exhibited a cyclical pattern in their foreign policy. Despite employing the concept of hedging, following the end of World War Two, ‘[Türkiye] joined the Western bloc against the Soviet Union’ (Coskun, 2023). This marked a shift in Turkish geopolitical strategy from a doctrine of hedging and maintaining relations with all global powers to adamantly denouncing their ties with the Soviet Union by joining the Western camp. Despite this, Turkish geopolitical actions on the global stage seemed to become increasingly more typical of a middle power. Türkiye’s significant participation in the Korean war, a theoretically multilateral UN endeavour ‘was a demonstration of good international citizenship, which in return opened the door for Turkey’s accession to the North Atlantic Treaty Organization (NATO) (Coskun, 2023).’ This illustrates Türkiye’s significant step to middlepowermanship as Coskun notes that ‘Turkey’s relative autonomy period of formation years came to an end with Turkey’s NATO membership’ (Coskun, 2023).

However, towards the late 1990s and early 2000s we observe Türkiye’s foreign policy becoming more unilaterally orientated and perhaps even pugnacious particularly in the regional geopolitical landscape. This turning point was marked by the invasion of Northern Cyprus by Turkish forces in 1974. Not only did this event inflame a centuries old conflict with historical rivals Greece because the ordeal ‘led to the end of the Turkish-Greek friendship that was in place throughout the 1950s (Coskun, 2023).’ The declaration of the Turkish Republic of Northern Cyprus further exacerbated the deterioration of Turkish-Greek relations as well as ostracising Türkiye from the US who imposed arms embargoes on Türkiye. This was followed by further exhibitions of unilateral military actions such as the numerous incursions into Northern Syria and direct military involvement in the Libyan Civil War (although they were backed by the UN). This, as previously explained, resembles a cyclical timeline of geopolitical strategy. In the early years of the Republic, Türkiye employed hedging in order to maximise economic gain while

mitigating the risks of being dragged into another conflict. Türkiye then began to shift to the Western camp of the global order in the subsequent decades which also led to Türkiye becoming more involved in multilateral initiatives. However, the major turning point in Turkish foreign policy occurred following the invasion of Northern Cyprus following which Türkiye repositioned its geopolitical strategy into one which provided autonomy and reinstated hedging practices.

Comparative Analysis

The geopolitical strategies of Türkiye and Azerbaijan, while possibly dissimilar in terms of specific foreign policy, both rely on hedging and not fully committing to one of the belligerents, both in an economic and political sense. Hedging helps to offset the regional instability instigated by an exogenous shock by balancing foreign policy in the interest of both sides. This strategy allows Türkiye and Azerbaijan to enjoy all the benefits of maintaining positive relations with both sides, such as trade, military and economic cooperation while avoiding confrontation and antagonism. Hedging greatly limits the risk of being forcibly dragged into an external conflict as a result of international pressure and potential collective defense clauses such as NATO's Article 5 (The North Atlantic Treaty", NATO, 2025). The political stability achieved through hedging reassures the public that the exogenous shock is a distant issue which will not affect their nation. This confidence is also replicated on a macroeconomic level. Consumers and firms maintain confidence, meaning that consumption and investment behavior remain stable. Without hedging, both consumers and firms may become apprehensive of external shocks disrupting their daily lives. As a result, consumers and firms will intensify their precautionary economic measures, such as heightened savings and low-risk investments, *ceteris paribus* leading to a decline in consumption and a downturn in stock prices, thus causing a fall in economic output. Moreover, disruptions of trade can lead to increased costs for the average consumer. The spike in the price of crude oil following the Russian invasion of Ukraine is a case in point, as it had a particularly pronounced effect on Western European consumers. Indeed, in the period leading up to the invasion and the weeks following, Brent Crude had increased by over 35% (Brent crude oil," Trading Economics, 2025). This rise in crude prices led to surging production costs, which were inevitably passed on to Western consumers. This development likely informed the geopolitical strategies of Türkiye and Azerbaijan following the commencement of the Russian invasion of Ukraine.

Azerbaijan has engaged in this delicate balancing act by continuing economic cooperation with Russia particularly in the development of natural resource infrastructure. Azerbaijan has continued being actively engaged in large-scale trade link initiatives such as the International North-South Transport Corridor, a 7200 km stretch of multi-mode routes designed for moving freight by ship, rail and road (International North-South Transport Corridor (INSTC)", Asia Regional Integration Centre, 2025). The major players of this initiative include Azerbaijan, Russia, Iran, India and other, primarily Central Asian, countries within Russia's geopolitical sphere of influence. Bilateral trade between Azerbaijan and Russia continued to

grow with Russia being Azerbaijan's largest import partner in 2022 making up 18.81% of the share (OECD World, 2025). Russia and Azerbaijan have also engaged in numerous joint ventures such as a joint venture established in 2022 to export fresh fruit and vegetables. Azerbaijan and Russia have also strengthened ties through joint investment into technology and artificial intelligence. For example, in 2023, Azerbaijan's Ministry of Digital Development and Transport signed an agreement with Russian companies specializing in cybersecurity and blockchain (Azerbaijan, Russia sign cybersecurity cooperation agreement", Interfax, 2025). Moreover, established Russian cybersecurity firms such as Kaspersky Lab have been setting up cybersecurity initiatives in Azerbaijan. These cases illustrate how tightly interwoven Russian and Azerbaijan industries regarding energy, transport and tech have become following the Russian invasion of Ukraine.

While on one hand the Azerbaijani diplomat acknowledges that "Azerbaijan has historical political, economic, trade, and cultural-humanitarian ties with Russia" (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025) and "strives for good neighborly relations with its immediate neighbors regardless of their foreign policy agenda (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025)", on the other hand, Azerbaijan has been rapidly developing both economic and political ties with the EU who are seeking diversification in energy partners following the imposition of trade sanctions on Russia. This was stated in the interview as the Azerbaijani Diplomat explained that the conflict had "highlighted the vulnerabilities of Europe's energy dependency, leading to greater emphasis on alternative energy services (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025)." Therefore, on July 18th 2022, the European Commission and the Republic of Azerbaijan signed a Memorandum of Understanding which sought diversification of energy imports away from Russia. The deal mentioned an increase in the supply of natural gas from Azerbaijan's reserves to various EU nations by committing to doubling the capacity of the Southern Gas Corridor to deliver at least 20 billion cubic metres to the EU annually by 2027 (EU and Azerbaijan enhance bilateral relations, including energy cooperation', European Commission, 2025). Although no further formal agreements have been signed as of January 2025, Azerbaijan has been continuing to increase natural gas exports and has emerged as the fourth largest supplier of piped gas to the EU. The interviewee further mentioned how "with the instability in Ukraine, the Middle Corridor [A trade route stretching from China to Eastern-European countries that passes through Azerbaijan] has gained increasing relevance, offering a reliable and secure route" (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025). Azerbaijan has been able to capitalize on this in order to "realize regional transport and communications projects" (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025). This demonstrates how Azerbaijan's geopolitical posture of hedging allowed it to capitalize on the EU's scrambling efforts in search of a new energy trading partner following Russia's invasion of Ukraine and according to the interviewee, Azerbaijan intends to "intensify cooperation" by

“[realising] concrete projects in this regard” (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025). Moreover, Azerbaijan has further strengthened their already amicable relations with Ukraine by having “actively participated in Ukraine’s recovery and humanitarian initiatives”(Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025).

Azerbaijan has also performed as a crucial intermediary between the Western bloc and Russia, albeit predominantly through multilateral campaigns unlike Türkiye. The Azerbaijani diplomat being interviewed commented that ‘Previously, Azerbaijan hosted important meetings at the high levels, including interactions between Russia and NATO, as well as the US and Russia (Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025).’ and that ‘Azerbaijan’s peace efforts concerning the conflict have primarily been in the context of multilateral platforms’(Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 2025). This demonstrates a quintessential characteristic of a middle power: leveraging the authority of large multilateral organisations in order to achieve a diplomatic resolution. This raises a further point over the customary traits of middle powers. While both Türkiye and Azerbaijan have attempted to act as an intermediary between Russia and Ukraine, Azerbaijan has sought a multilateral orientated approach while Türkiye has opted to take the initiative and be the sole leader of staging diplomatic efforts. This perhaps highlights the differences within the subset of middle powers. Türkiye may have adopted a unilateral approach to its middleman role due to its relatively greater strength and influence. However, it is also possible that it aimed to enhance its political leverage and global standing by presenting itself as a dedicated participant in the multilateral system.

Türkiye has followed a similarly pragmatic approach to hedging relations with Russia and the Western camp. Similarly to Azerbaijan, Türkiye continues to cooperate with Russia in the development and operation of energy transport links, one of which being the TurkStream pipeline which originates near the Russian city of Anapa and spans 930 km of offshore territory across the black sea to North Eastern Türkiye from which natural gas is further distributed to South-Eastern and Central European countries. Significantly, Russia’s incentive in this project was diverting the flow of natural gas from transiting through Ukraine and the Trans Balkan Pipeline System. This indicates a clear illustration of Türkiye directly benefiting from the conflict in Ukraine as Russia shifts energy infrastructure routes away from Ukraine. This is made even more explicit as Turkish President Recep Tayyip Erdoğan and Vladimir Putin made plans for Türkiye to become a European energy hub, which will continue to economically benefit Türkiye by attracting further infrastructure investment, job creation and even providing Türkiye with greater political leverage on the world stage. This example illustrates the successes of Azerbaijan and Türkiye’s foreign policy amidst the Russian invasion of Ukraine. A neutral approach to the conflict has resulted in both countries experiencing increased economic cooperation with Russia through further energy infrastructure deals.

Unlike Azerbaijan, Türkiye has also leveraged its unique middle power status to strengthen relations with Ukraine, despite continuing to cooperate with Russia in economic deals. Firstly, President Erdogan has publicly proclaimed Türkiye's support for Ukraine and their efforts in defying Russian military aggression. During a bilateral meeting with Zelensky almost 3 years after Russia's initial invasion of Ukraine, Erdogan reaffirmed his support for 'Ukraine's complete territorial integrity, sovereignty, and independence (Turkey 'ideal host' for peace talks on Russia-Ukraine war, Erdogan says", The Kyiv Independent, 2025).' as well as stating that 'this war must end', highlighting the determinants of long term conflict, even though Türkiye may be an economic beneficiary due to its hedging practices. Indeed, Türkiye has been widely commended by Western and Ukrainian media for the sale and donation of military weapons to the Ukrainian Armed Forces. The predominant focal point of military equipment provided to Ukraine was the indigenously produced Bayraktar TB2 drone, the now infamous unmanned combat aerial vehicle (UCAV) used extensively by Ukrainian forces to repel Russian armoured columns in the early stages of the war. Ukraine's success on the battlefield indirectly bolstered Türkiye's economy, as video footage showcasing the precision and lethality of drone strikes served as an invaluable form of global marketing for Baykar, the manufacturer of the drones, which subsequently recorded export revenues of \$1.8 billion in the following year. Türkiye also further reaffirmed its support to Ukraine and NATO by closing the Bosphorus Strait to Russian Warships, severely limiting Russian naval capabilities in the Black Sea. Furthermore, Türkiye has leveraged its position as a middleman between the two conflicting blocs in order to achieve its regional foreign policy. Dissimilarly from Azerbaijan, Türkiye's membership of NATO gives it increased political leverage during times of regional instability caused by exogenous shocks. Türkiye repeatedly prevented the ratification of Sweden and Finland into NATO, accusing the two Nordic countries of allowing internationally recognised terrorist groups such as the PKK to operate within their territory. Amid growing concerns over the potential escalation of the conflict into North-Eastern Europe, Türkiye dismissed Sweden and Finland's apprehensions regarding their national sovereignty, a stance that strained their relations with NATO. However, while Türkiye does somewhat break the convention of middle powers being very invested in multilateralism, their efforts in staging diplomatic talks between Russia and Ukraine exemplifies a key attribute of middlepowermanship. Türkiye acted as a rational intermediary, notably by hosting peace talks in Istanbul. Although these diplomatic efforts did not come to fruition in reaching a ceasefire, the ordeal demonstrated Türkiye's role as a responsible regional power, further enhancing its diplomatic standing with all parties engaged in the conflict. While Azerbaijan followed a similar policy of neutrality, it likely lacked the geopolitical influence to stage diplomatic talks. This key difference illustrates the breadth of global standing which the term middle power encompasses, perhaps even implying it is an overly simplistic and arbitrary label for countries not in the highest or lowest echelons of the global geopolitical order.

Conclusion

Although the cases of Azerbaijan and Türkiye illustrate how middle powers may use a hedging strategy in response to external shocks, the notion that such shocks inherently benefit a nation seeking to enhance its geopolitical standing is misleading. An exogenous shock such as a violent conflict undoubtedly encapsulates numerous devastating social impacts on a number of populations. Not only do the calamitous consequences of war on regular citizens raise ethical questions over the practice of adjusting foreign policy to benefit from conflict, a conflict involving a large power typically leads to longstanding regional instability and soured relations, creating a less cohesive global order. A world which refrains from diplomatic and cordial dialogue is detrimental to the social, political and economic wellbeing of individual nations. The past few decades have experienced increasing polarisation within the global community, partly caused by the rise of emerging middle powers. Established multilateral organisations such as the UN are facing increasingly inflamed criticisms due to their lack of authority. Despite the widely-accepted global dominance of the USA, the current global landscape has never been more ideal for emerging middle powers looking to expand their geopolitical influence. However, the long-term implications of relying on hedging strategies may accompany significant adverse consequences for global geopolitics. Repeatedly leveraging hedging practices for national gain during exogenous shocks—particularly those that threaten the sovereignty of other states—risks establishing a dangerous precedent that emboldens the imperialistic ambitions of certain nations. It remains imperative that states which instigate conflicts or invasions are unequivocally condemned and held accountable, as this is essential to deterring further territorial expansion and imperialist sentiment.

While this paper has been beneficial in consolidating the understanding of middle powers and their geopolitical strategy, the cases employed are Eurocentric and are likely unrepresentative of the global subset of middle powers. Further questions may be raised about the geopolitical role of middle powers in East Asia, amidst China's territorial regional foreign policy, particularly in the South China Sea. An escalation into a widespread conflict between the USA and China in this region would likely precipitate far-reaching global consequences. How would middle powers react? As stated previously, middle power theory is still a novel subfield in international relations, and one which rapidly transforms and is shaped by the incessantly evolving geopolitical landscape. Will the emergence of more incisive and authoritative middle powers lead to the inefficacy and collapse of multilateral institutions, or will the outspokenness of middle powers provide a more balanced world view - thus enhancing multilateralism and global cooperation?

Works Cited

- Carr, Andrew. "Is Australia a middle power? A systemic impact approach." In the Australian Journal of International Affairs. 2014
<https://www.tandfonline.com/doi/abs/10.1080/10357718.2013.840264>
- Cooper, Andrew F., Higgott, Richard. A., & Nossal, Kim. R. "Relocating middle powers: Australia and Canada in a changing world order." UBC Press, 1993
<https://www.ubcpress.ca/relocating-middle-powers>
- Cooper, Andrew F. (Ed.), "Niche diplomacy: Middle powers after the Cold War". 1997
- Coşkun, B.B. (2023). One Hundred Years of Turkish Foreign Policy in Eastern Mediterranean (1923–2023):
A Neoclassical Realist Analysis of Turkey as a Middle Power. In: Özkeçeci-Taner, B., Akgül Açıkmış, S. (eds) One Hundred Years of Turkish Foreign Policy (1923-2023). Global Foreign Policy Studies. Palgrave Macmillan, Cham.
https://doi.org/10.1007/978-3-031-35859-3_7
- Laurence, Marion. "Middle Powers." Oxford Research Encyclopedia of International Studies. 18 Oct. 2023; Accessed 9 Feb. 2025.
<https://oxfordre.com/internationalstudies/view/10.1093/acrefore/9780190846626.001.0001/acrefore-9780190846626-e-783>
- Senior Diplomat, Ministry of Foreign Affairs of the Republic of Azerbaijan, written interview (email), 11/02/25
- Jones, Dorian. (October 20, 2022). Erdogan Agrees to Putin's Plan for Turkey to Be Russian Gas Hub. VOA.
<https://www.voanews.com/a/erdogan-agrees-to-putin-s-plan-for-turkey-to-be-russian-gas-hub/6798604.html>
- Budapest Business Journal (November 21, 2019). Turkstream project enters final stage.
Retrieved from
<https://bbj.hu/politics/foreign-affairs/intl-relations/turkstream-project-enters-final-stage/>
- EU Directorate General for Neighbourhood and Enlargement Negotiations. (March 4, 2024). EU steps up renewable energy cooperation with Azerbaijan. Retrieved from
https://enlargement.ec.europa.eu/news/eu-steps-renewable-energy-cooperation-azerbaijan-2024-03-04_en
- Trading Economics. Brent crude oil. Retrieved from
<https://tradingeconomics.com/commodity/brent-crude-oil>
- Interfax. (June 24, 2022). Azerbaijan, Russia sign cybersecurity cooperation agreement.
Retrieved from <https://interfax.com/newsroom/top-stories/80608/>
- Axe, David. (September 03 2023). Ukraine's TB-2 Drones Are Back In Action. That's An Ominous Sign For Russia. Retrieved from
<https://www.forbes.com/sites/davidaxe/2023/09/03/ukraines-tb-2-drones-are-back-in-action-thats-an-ominous-sign-for-russia/>

TRTWorld and Agencies. Türkiye's TAI, Baykar rank in the world's top 50 aviation firms.
Retrieved from
<https://www.trtworld.com/turkiye/turkiyes-tai-baykar-rank-in-worlds-top-50-aviation-firms-18230968>

Reuters. (April 05 2023). Why are Turkey and Hungary against Sweden joining NATO?.
Retrieved from
<https://www.reuters.com/world/europe/after-finland-joins-nato-why-is-turkey-making-sweden-wait-2023-04-04/>

Aris, Ben. (April 17 2024). Fresh evidence suggests that the April 2022 Istanbul peace deal to end the war in Ukraine was stillborn. Retrieved from
<https://www.intellinews.com/fresh-evidence-suggests-that-the-april-2022-istanbul-peace-deal-to-end-the-war-in-ukraine-was-stillborn-321468/>

“GDP, PPP (current international \$)”. World Bank Group. Accessed February 21, 2025,
<https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.CD?end=201>

“Nominal gross domestic product” (GDP). OECD. Accessed February 21, 2025,
<https://www.oecd.org/en/data/indicators/nominal-gross-domestic-product-gdp.html>

“2025 Military Strength Ranking”. Global Firepower. Accessed February 21, 2025.
<https://www.globalfirepower.com/countries-listing.php>

“The North Atlantic Treaty”. NATO. Accessed February 20, 2025.
https://www.nato.int/cps/eu/natohq/official_texts_17120.htm

“Turkey 'ideal host' for peace talks on Russia-Ukraine war, Erdogan says”. The Kyiv Independent. Accessed February 20, 2025.
<https://kyivindependent.com/turkey-ideal-host-for-peace-talks-on-russia-ukraine-war-erdogan-tells-zelensky/>

“International North-South Transport Corridor (INSTC)”. Asia Regional Integration Centre. Accessed March 01, 2025,
<https://aric.adb.org/initiative/international-north-south-transport-corridor>

OECD World, Accessed March 01, 2025,
<https://oec.world/en/profile/bilateral-country/rus/partner/aze>

Re-Interpreting the Sent Down Youth: Heidegger, Arendt, and the Vicious Cycle of Worldlessness in the Cultural Revolution By En (Chloe) Ke

Introduction

The Cultural Revolution in China (1966-1976) was a period of extreme social upheaval marked by political purges, mass re-education, and societal restructuring. Amidst the turmoil, one of the key elements of this transformation was the “sent-down youth” movement, in which urban educated youths were forcibly relocated to rural areas to be “re-educated” by farmers. This was a part of the Chinese Communist Party (CCP)’s effort to close the perceived ideological gap between the urban elite and the rural peasantry (Baum, 2009). These youths, removed from their familiar urban environments and forced into unfamiliar rural contexts, faced profound challenges of isolation, identity loss, and disconnection.

This essay employs the tools of existential philosophy to explore the predicament of the sent-down youth and the relationship between their experience and the totalitarian politics of the CCP (Baehr, 2010). Specifically, the essay uses a philosophical framework inspired by Martin Heidegger’s concept of *worldliness* and the related concept of *worldlessness* developed by Hannah Arendt. This philosophical approach offers a unique angle of analysis, bringing to light the cyclical nature of worldlessness and totalitarianism while connecting the experiences of the sent-down youth to more recent developments in China’s political landscape. Concretely, the essay argues that the Cultural Revolution, and the “sent-down youth” program in particular, created a *vicious cycle of worldlessness*, where totalitarianism entrenched itself through conditions of collective alienation.

The argument proceeds in three steps. The essay’s first section provides historical background and context concerning the Cultural Revolution and the sent-down youth movement, highlighting the importance of a philosophical account of these events. Next, the essay treats the concepts of worldliness and worldlessness in Heidegger and Arendt. Taken together, these concepts provide a philosophical framework for analyzing the vicious cycle of worldlessness produced by the sent-down youth movement. The third section makes a case for this vicious cycle, arguing that the sense of isolation among the sent-down youth, exacerbated by their forced relocation, made it easier for the regime to impose ideological conformity. This, in turn, deepened the youths’ sense of alienation, reinforcing the regime’s control and perpetuating totalitarianism.

Historical context of the Cultural Revolution

The “sent-down youth” movement, also known as the “Up to the Mountains and Down to the Countryside Movement” (上山下乡运动), was a major social campaign launched during the later stages of the Cultural Revolution in China, particularly between 1968 and 1978. Initiated by Chairman Mao Zedong, this movement was intended to address the political and economic instability that had emerged from the social upheavals of the Cultural Revolution, and also to reshape the ideological character of the country’s urban youth (Macfarquhar M & Schoenhals,

2009). The Cultural Revolution, which began in 1966, was a period of radical political and social reform in China aimed at enforcing communist orthodoxy and eliminating traditional and capitalist elements from society. Urban intellectuals, professionals, and students were among the primary targets of political persecution, with the Communist Party seeking to “re-educate” them in the virtues of peasant labor and rural life. Mao believed that sending millions of young people, particularly those from cities, to live and work in rural areas would help to bridge the perceived ideological divide between urban and rural populations, cementing values deemed important by the Communist Party.

The “sent-down youth” movement began in earnest after Mao’s declaration of December 22, 1968, published on *Renmin Ribao*, that “It is necessary for educated youth to go to the countryside to receive re-education from the poor and lower-middle peasants”(Ribao, 2019). Between 1968 and 1979, over 17 million urban youths, often high school and college students, were sent to remote rural areas where they were required to live and work alongside rural peasants (King et. al., 2010). The official purpose of this campaign was to expose these youths to the “real China” and teach them the virtues of manual labor, which was considered more ideologically pure than intellectual work. However, many sent-down youths experienced extreme hardships, including harsh living conditions, inadequate food and medical care, and a profound sense of disconnection from their urban roots. “The environment was simply too bleak,” wrote Chinese liberal historian Qin Hui, a former sent-down youth. “Looking back now, how did I hang on for so long without giving up? Because I had ‘ideals’. But when you think about it, there was no choice: in that kind of situation, what could you pursue other than that illusory ‘ideal’?” (Ownby, 2016) Of the millions of sent-down youths, many felt alienated. Cut off from their families and former lives in the cities, they struggled with the labor-intensive lifestyle, lack of infrastructure, and limited opportunities for personal or intellectual growth. While the movement was intended to build solidarity between urban youth and rural populations, it instead led to widespread feelings of disillusionment(Bonnin, 2013).

The return of these youths to their urban homes in the late 1970s further highlighted the failure of the program. Many came back disenchanted, having spent years performing menial labor without educational or career prospects. As historian Qin Hui notes, “people give in or give up, and many Zhiqing (sent-down youths) became thoroughly cynical and pessimistic, and stayed that way for the rest of their lives (Ownby, 2013).” Eventually, their experiences in the countryside shaped a generation marked by lost opportunities and lingering discontent. Despite its official aims, the sent-down youth movement became a symbol of the harsh social and political experiments of the Cultural Revolution.

Before concluding this section, it is worth noting that the latter half of the Cultural Revolution has been widely studied across various disciplines: historians have traced the underlying motivations of Party actors and civilian participants, economists have meticulously weighed the financial costs, and sociologists have explored the societal shifts resulting from this upheaval (Yu et. al., 2015). However, a philosophical account of this world-historical sequence of events has not been proposed before. This absence of philosophical attention is unfortunate: as

the following sections will show, bringing the tools of philosophy to explore the “sent-down youth” movement helps us make sense of the failures of this momentous experiment in social and ideological engineering. What exactly were the human costs of this project, and how did the experience of sent-down youth shape totalitarian politics in the Cultural Revolution and beyond? The next section lays out the elements of a philosophical framework that will enable us to approach this question in a new light.

Philosophical framework

In order to track the vicious cycle of worldlessness experienced by the sent-down youth, it is necessary to begin with Martin Heidegger’s concept of *worldliness*, one of the foundational concepts of his early philosophy. In “*Being and Time*”, Heidegger argues that worldliness—or “being-in-the-world”—is a necessary ontological feature of human existence. Rather than being ‘subjects’ detached from ‘objective’ reality, human beings are “always already” entwined with the world in which they appear (Heidegger,2010). Yet, the ‘world’, according to Heidegger, is not merely the physical combination of discrete objects; rather, it is an intricate web of relations shot through with significance, a meaningful “structure which is primordially and constantly whole (Heidegger,2010).” Put differently, the ‘world’ is a holistic context of meaning in which humans find themselves embedded: it serves as the bedrock for, and at the same time encompasses, all manner of interpersonal relations, social practices, and the phenomena of culture (Dahlstrom, 2025).

According to Heidegger, human beings are always equipped with a pre-theoretical—or in his words, “pre-ontological”—understanding of their surrounding world (Heidegger,2010). This means that the world is not an alien environment, but one in which we are ordinarily ‘at home’. The things that surround us are, for the most part, familiar things, readily available to us, and we instinctively know how to interact with them. Heidegger uses a carpenter in his workshop as a prime example to illustrate the familiar experience of *being-in-the-world*. In this scenario, the carpenter’s tools, such as the hammer, are not perceived as isolated objects; rather, they are seamlessly integrated into the activity of hammering. Heidegger goes on to emphasize that the carpenter does not consciously reflect on the hammer while using it—instead, its usefulness and function are so embedded in the task that the tool becomes transparent to the carpenter’s activity: “the less we just stare at the thing called hammer, the more we take hold of it and use it (Heidegger,2010).” Heidegger suggests the carpenter’s experience is emblematic of human experience, where historically-bound cultural practices and norms impart shared meanings on the things that surround us. The comfort and integration afforded by the familiar background of our ‘world’ enable us to carry out our personal and communal projects with ease.

Several decades after the publication of *Being and Time*, the concept of worldliness was taken up and transformed by one of Heidegger’s students, Hannah Arendt. Like Heidegger, Arendt suggests that the ‘world’ consists in the shared, man-made reality that exists between people and connects them through common objects, spaces, and ideas. Yet, for Arendt, this shared world is linked to politics and is coextensive with the *public realm*, where human

activities such as speech, action, and interaction take place, and through which individuals relate to one another as political agents and equals. As Arendt writes, “the public realm ... is the common world,” which “comes into being wherever men are together in the manner of speech and action (Arendt, 2018).” In Arendt’s view, then, the world is more than just a space of familiar objects whose meaning is handed down to us historically; it is a shared reality continually created and enacted through collective endeavor, speech, and storytelling. To live together in the world, Arendt states, “means essentially that a world of things is between those who have it in common, as a table is located between those who sit around it”(Arendt, 2018).

From this, we can see that although related, Heidegger’s and Arendt’s notions of ‘world’ differ fundamentally. For Heidegger, *worldliness* is rooted in an existential framework, meaning that the world is an inseparable aspect of “*Dasein*,” or human existence (Arendt, 2018). Put differently, worldliness is not something that can be lost; it’s a constitutive dimension of human life. To be human is to “always already” be in a world, navigating and interpreting one’s surroundings through familiar structures and relationships (Arendt, 2018). In sharp contrast to this ontological approach, Arendt treats the world as a contingent and fragile reality that humans create through shared relationships and the interchange of diverse views and opinions; for her, “the reality of the public realm relies on the simultaneous presence of innumerable perspectives (Arendt, 2018).” The world, in Arendt’s view, is not an ontological given but a product of human action, which is created in the constantly varying fabric of political and social life. As such, the world can easily be shattered or destroyed.

Crucially for the purposes of this essay, Arendt contrasts the idea of a shared, public world with the condition of *worldlessness*, in which individuals are deprived of the shared spaces and relationships that give their lives public meaning. According to Arendt, such alienation often takes root within totalitarian regimes, where heterogeneity is discouraged and even persecuted while the erosion of public space eliminates the possibility for meaningful communication, leaving individuals isolated and disconnected. In a *worldless* society, individuals are not connected through a shared sense of belonging; they are left to exist in private isolation, unable to form collective bonds or generate political agency (Arendt, 2018). In part, this owes to the erasure of unique perspectives: totalitarian regimes give rise to a community where homogeneity eradicates “plurality,” which, for Arendt, is the essential bedrock of a public realm and the “*conditio per quam* of all political life (Arendt, 2018).” “The destruction of the common world,” Arendt suggests, “is usually preceded by the destruction of the many aspects in which it presents itself to human plurality (Arendt, 2018)” In short, worldlessness is a profound and tragic consequence of the desiccation of the public sphere; the collapse of shared spaces for speech and action leaves individuals disconnected from the collective structures that give their lives meaning.

The Vicious Cycle of Worldlessness

Equipped with the philosophical framework sketched above, we are prepared to understand how the Cultural Revolution, and the sent-down youth movement in particular,

created a condition of worldlessness that emerged from, and then perpetuated, totalitarianism.

In *Being and Time*, Heidegger explains that human beings are always “being-in-the-world,” inherently embedded within a familiar context of meaning that gives their lives structure and purpose. Nevertheless, as urban youths were forcibly relocated to rural areas, they were abruptly removed from the familiar world they knew. Much like a carpenter suddenly robbed of his tools because he has been placed in a completely foreign environment, the sent-down youths found themselves in a completely alien world—the social roles, relationships, and norms that had previously given their lives meaning in the city no longer applied in the countryside. As a result, the youths faced a profound disruption of their sense of belonging and purpose. In this new and alien rural environment, they struggled to find meaning, leading to an overwhelming sense of isolation and purposelessness.

While Heidegger helps us grasp the disorientation brought about by the youths’ removal from their familiar environment, Arendt helps us understand the broader political and social isolation that resulted. As mentioned above, Arendt holds that totalitarian regimes thrive on the destruction of public “space between people” where individuals come together, share diverse perspectives, and form collective bonds (Arendt, 2018). In the case of the sent-down youths, they were not only alienated by the unfamiliar rural environment but also isolated by the regime’s suppression of plurality. Jung Chang, in her memoir *Wild Swans*, recalls “The need to obtain authorization for an unspecified ‘anything’ was to become a fundamental element in Chinese Communist rule. It also meant that people learned not to take any action on their own initiative (Chang, 2015). Under Mao’s regime, ideological conformity was imposed, and any dissent or deviation from the party’s narrative was brutally suppressed (Phillips, 2016).

As Arendt argues, the loss of diversification subsequently leads to a privatization of human beings, “that is, they [are] deprived of seeing and hearing others, of being seen and being heard by them”(Arendt, 2018). As in the case of the cultural revolution, the regime's suppression of plurality forced the sent-down youth to internalize their thoughts and suppress any expression of individuality. This condition of privatization persisted and dwelled on the youths as a permanent scar, as an anonymous former sent-down youth phrases: “To this day they share a habit that I have noticed in a lot of Chinese individuals of a certain age group. When saying something even slightly confidential or personal, they cover their mouth with their hands to prevent any outside observers from reading their lips”(Hays, 2021). Under the totalitarian regime of the Cultural Revolution, the vital prerequisite for a public realm—namely, what Arendt calls the “space of appearance” where human beings share their diverse perspectives as equals—collapsed, and the youths were trapped in a condition where they found their individualities a misfit for their new life (Arendt, 2018). Their ‘world,’ in the Arendtian sense, was thus shattered, insofar as they were deprived of the social and political interactions that create a common world.

This combination of Heideggerian disorientation and Arendtian worldlessness played directly into the hands of the totalitarian regime. As the youths became more alienated and disconnected from one another, they were less able to resist the ideological indoctrination

imposed upon them. Just as Qin Hui queried: “what could you pursue other than that illusory ‘ideal’?” (Ownby, 2016) Such psychology speaks to the existential void faced by many sent-down youths—stripped of their previous worlds, they clung to ideals handed down by the regime, even as these ideals failed to provide real meaning or belonging, furthering the disorientation and estrangement experienced by the sent-down youths.

Qin Hui’s account clarifies our understanding of the experience of ideological control. He notes that for many sent-down youths, holding onto “illusory ideals” became a form of survival. Even when these ideals were ultimately “illusory,” they provided a kind of “spiritual pillar” that helped youths make sense of their situation and even became a subpar substitute or imitation of what “being at home” was once like. Through abiding by the totalitarian doctrine, the youths sought to find a public space where their voices were echoed, despite such interactions being out of blind devotion to an enforced political propaganda. As Qin Hui explains, “ideology itself was playing the role of a drug, and from a certain perspective, people need drugs, particularly when they find themselves in hopeless or desperate situations (Ownby, 2016).”

The hopelessness and desperation experienced in the context of the sent-down youth movement—functions of the worldlessness described above—made it difficult for sent-down youths to form meaningful connections. Furthermore, without a shared public realm, they lacked the capacity to challenge the ideological narrative imposed on them. In this sense, worldlessness did not merely result from totalitarian control; it actively contributed to its perpetuation. Cut off from their familiar urban environments and forced into ideological isolation, the youths found themselves trapped in a system that left no room for critical thought or dissent. The regime’s efforts to re-educate the youth by severing them from their urban roots led to a vicious cycle: worldlessness fueled totalitarianism, and totalitarianism deepened worldlessness.

Conclusion

The sent-down youth movement of the Cultural Revolution was a profound social experiment that left a generation of urban intellectuals disoriented, alienated, and disconnected from the world they once knew. Using Heidegger’s concept of worldliness, we understand that these youths, uprooted from their familiar urban environments, faced an existential void, as their tools of meaning-making—social roles, cultural practices, and relationships—were stripped away in the unfamiliar rural context. Such alienation was compounded by the totalitarian regime’s control, which, as Arendt’s concept of worldlessness illustrates, systematically destroyed the public spaces where individuals could come together, share diverse perspectives, and create meaning collectively.

As this essay has shown, the combination of Heideggerian disorientation and Arendtian worldlessness provided fertile ground for the perpetuation of totalitarianism. The sent-down youths, disconnected from both the urban and rural worlds, were unable to resist the regime’s ideological control, which only deepened their alienation. This vicious cycle of worldlessness, the essay suggests, helps us understand the persistence and entrenchment of totalitarianism. In the end, the movement that aimed to create ideological unity instead created a generation scarred

by lost opportunities, disillusionment, and a life-long discontent—a cautionary tale of the devastating human costs of authoritarian social and political experiments.

Works Cited

- Arendt, Hannah. *The Human Condition*. Chicago: The University of Chicago Press, 2018.
- Baehr, Peter. "China the Anomaly." *European Journal of Political Theory* 9, no. 3 (July 2010): 267–86. <https://doi.org/10.1177/1474885110363981>.
- Bonnin, Michel. *The Lost Generation: The Rustication of China's Educated Youth (1968-1980)*. Hong Kong: Chinese University Press, Cop, 2013.
- Chang, Jung. *Wild Swans: Three Daughters of China*. Tullamarine, Victoria: Bolinda Audio, 2015.
- Chen, Yi, Ziying Fan, Xiaomin Gu, and Li-An Zhou. "Arrival of Young Talent: The Send-down Movement and Rural Education in China." *American Economic Review* 110, no. 11 (November 1, 2020): 3393–3430. <https://doi.org/10.1257/aer.20191414>.
- Dahlstrom, Daniel. "World (Welt)." *Cambridge University Press EBooks*, April 17, 2021, 822–29. <https://doi.org/10.1017/9780511843778.220>.
- Hays, Jeffrey. "Sent-down Youth and back to the Countryside Movement of the Cultural Revolution." [factsanddetails.com](https://factsanddetails.com/china/cat2/sub6/entry-5533.html), August 2021. <https://factsanddetails.com/china/cat2/sub6/entry-5533.html>.
- Heidegger, Martin, Dennis J Schmidt, and Joan Stambaugh. *Being and Time*. Albany: State University Of New York Press, 2010.
- Honig, Emily, and Xiaojian Zhao. "Farewell to the Huangpu River." *Cambridge University Press EBooks*, August 30, 2019, 18–39. <https://doi.org/10.1017/9781108595728.002>.
- . "Sent-down Youth and Rural Economic Development in Maoist China." *The China Quarterly* 222 (March 31, 2015): 499–521. <https://doi.org/10.1017/s0305741015000363>.
- Kattago, Siobhan. "Hannah Arendt on the World," n.d. <https://grattoncourses.wordpress.com/wp-content/uploads/2017/08/siobhan-kattago-hannah-arendt-on-the-world-hannah-arendt-key-concepts.pdf>.
- King, Richard, Ralph C Croizier, Scott Watson, and Sheng Tian Zheng. *Art in Turmoil: The Chinese Cultural Revolution, 1966-76*. Vancouver: Ubc Press, 2010.
- MACFARQUHAR, Roderick, and Michael Schoenhals. *Mao's Last Revolution*. Harvard University Press, 2009.
- Ownby, David. "Qin Hui on Sent-down Youth." *Reading the China Dream*, 2016. <https://www.readingthechinadream.com/qin-hui-on-sent-down-youth.html>.
- Phillips, Tom. "The Cultural Revolution: All You Need to Know about China's Political Convulsion." *The Guardian*, May 11, 2016. <https://www.theguardian.com/world/2016/may/11/the-cultural-revolution-50-years-on-all-you-need-to-know-about-chinas-political-convulsion>.
- Xie, Yu, Yang Jiang, and Emily Greenman. "Did Send-down Experience Benefit Youth? A Reevaluation of the Social Consequences of Forced Urban–Rural Migration during China's Cultural Revolution." *Social Science Research* 37, no. 2 (June 2008): 686–700. <https://doi.org/10.1016/j.ssresearch.2007.08.002>.

Cybersecurity Reimagined: Data-Driven Defenses for the Modern Digital Landscape

By Ahil Thendral

Abstract

The research paper delves into the critical realm of data security in the modern digital landscape. It begins by emphasizing the paramount importance of safeguarding data in an era where information is stored, processed, and transmitted across various platforms and networks. The paper explores existing methods of data security, covering fundamental techniques such as encryption, access control, authentication, data masking, backup strategies, network security, and endpoint security. Despite the prevalence of these methods, the paper highlights their insufficiency in addressing the evolving threat landscape. Traditional approaches are found to be static, reactive, and perimeter-focused, often struggling to combat advanced cyber threats and insider risks effectively. The paper underscores the need for innovative and comprehensive data security strategies to mitigate risks and protect sensitive information in an increasingly interconnected and complex digital environment.

Introduction

People all over the world are very worried about data security these days because they depend so much on technology for many things in their daily lives. A lot of personal photos, financial data, sensitive emails, and social media accounts are stored on home computers and sent over the internet. Even though many security measures are in place at home and while data is being sent, they are often not enough to keep up with new cyber threats. Users face many security risks at home, including malware attacks like viruses, ransomware, and spyware, as well as sophisticated phishing scams that try to trick people into giving away private information. A lot of people use bad security habits, like weak passwords and old software, which makes vulnerabilities even worse and leaves personal information open to being used by cybercriminals. Also, the rise of Internet of Things (IoT) devices has made it easier for cybercriminals to get to users' personal data. They do this by targeting these connected devices.

Data sent over the internet can be intercepted and decrypted, even though encryption protocols like HTTPS are meant to keep people from getting to it without permission. Plus, illegal downloading of music, movies, and other media online is very bad for data security. It hurts the integrity of intellectual property rights and costs the world economy billions of dollars every year. Not only does pirated content keep content creators from making money, it also puts consumers at risk because it often has malware and other harmful software. Companies face a direct threat to their profits and brand reputation from online piracy. This shows how important strong anti-piracy measures are. The decentralized nature of the internet and users' ability to remain anonymous have made it hard to fight online piracy through legal means and technological solutions. Even with these problems, dealing with the growing threats to data security needs new ideas that use cutting edge technologies like data science, AI, and ML.

Better cybersecurity practices could be made possible by these technologies, which let companies find, stop, and respond to cyber threats more quickly and effectively. In a digital world that is becoming more and more connected, organizations can strengthen their defenses, find strange activities, and keep sensitive data safe by using data analytics, AI algorithms, and machine learning (ML) models. Bringing together data science, AI, and ML through collaborative research and work between different fields is the key to solving the difficult problems of cybersecurity in the digital age.

Cybersecurity and Data Science

Cybersecurity

Cybersecurity is the process of keeping computer systems, networks, programs, and data safe from hacking, damage, theft, and other digital attacks. In today's interconnected world, where people, businesses, and governments depend on digital technologies for communication, commerce, and critical infrastructure, cybersecurity measures are very important.

Cyberattacks come in many forms, such as ransomware, phishing scams, malware infections, denial-of-service (DoS) attacks, and threats from inside the company. These attacks can cost money, compromise data, hurt your reputation, and even stop basic services from working, impacting individuals and corporations.

A multi-layered defense strategy is essential for effective cybersecurity (Whitman & Mattord, 2018). This includes network security measures like firewalls, intrusion detection systems (IDS), and virtual private networks (VPNs) to safeguard data transmitted over networks, such as the internet or internal networks, ensuring its privacy and security.

Endpoint security involves utilizing antivirus software, encryption, and access control to protect individual devices like computers, smartphones, and tablets from malware and unauthorized access. Identity and Access Management (IAM) practices are crucial for monitoring user identities, limiting their access to resources, and preventing unauthorized access to sensitive data. Data protection strategies, including encryption, data loss prevention (DLP), and backup solutions, ensure the safety and privacy of data during transmission and storage, while also adhering to privacy laws.

Application security focuses on safeguarding software programs and systems from bugs and attacks by implementing secure coding methods, conducting penetration tests, and addressing vulnerabilities. Additionally, security training and awareness initiatives educate employees and users about cybersecurity best practices, such as password protection and recognizing social engineering attempts. Incident response and disaster recovery plans are developed to effectively manage and recover from cybersecurity incidents, such as data breaches and system compromises, ensuring minimal disruption and damage.

Cybersecurity is important to keep digital assets safe, keep people trusting online services and systems, and lower the risks that cyber threats pose in today's digital world.

Data Science

Data science is an interdisciplinary field that involves extracting insights and knowledge from structured and unstructured data using scientific methods, algorithms, and tools. It includes various techniques and approaches from statistics, computer science, ML, and domain-specific knowledge to analyze complex datasets and solve data-driven problems.

According to Provost and Fawcett (2016), Data science is all about making sense of information. It starts with gathering data from different sources like databases, websites, and sensors, then cleaning it up and exploring to see what's there, called exploratory data analysis (EDA). Next, we find what matters most and build models to predict things based on the data. After testing to make sure everything works, we put our tools to work solving real problems. Finally, we keep an eye on things to ensure everything runs smoothly. By following these steps, data scientists unlock valuable insights that help businesses make better decisions and improve people's lives..

Data science finds applications across various domains and industries, including healthcare, finance, marketing, cybersecurity, retail, and manufacturing. It enables organizations to extract valuable insights from data, optimize processes, improve decision-making, and drive innovation. By leveraging data science techniques and technologies, businesses can gain a competitive edge, identify new opportunities, and address complex challenges in today's data-driven world.

Artificial Intelligence

Artificial Intelligence (AI) is a branch of computer science that focuses on creating intelligent machines capable of performing tasks that typically require human intelligence. These tasks include understanding natural language, recognizing patterns, learning from experience, making decisions, and solving problems. AI systems aim to replicate cognitive functions such as perception, reasoning, learning, and decision-making to automate tasks, assist human operators, and achieve specific goals.

According to Russell, and Norvig (2016), AI encompasses a wide range of techniques, algorithms, and approaches. ML algorithms enable computers to learn from data and improve their performance over time without explicit programming. Supervised learning, unsupervised learning, and reinforcement learning are common approaches in ML, utilized for tasks such as classification, regression, clustering, and reinforcement learning. Deep Learning, a subset of ML, involves neural networks with multiple layers (deep architectures) capable of learning complex representations from raw data, achieving remarkable success in areas like image recognition, speech recognition, natural language processing, and autonomous driving. Natural Language Processing (NLP) focuses on enabling computers to understand, interpret, and generate human language in a meaningful way, utilized in applications such as language translation, sentiment analysis, chatbots, and text summarization. Computer Vision enables computers to interpret and understand visual information from images or videos, facilitating applications like object detection, image classification, facial recognition, and medical image analysis. Robotics combines AI with mechanical systems to create autonomous or semi-autonomous robots capable

of performing physical tasks in various environments, ranging from industrial automation and manufacturing to service robots, healthcare robots, and autonomous vehicles. Expert Systems are AI systems designed to mimic the problem-solving abilities of human experts in specific domains, utilizing knowledge representation, inference engines, and rule-based reasoning to make decisions and provide recommendations in areas such as medicine, finance, and engineering. Reinforcement Learning involves training agents to interact with an environment and learn optimal behavior through trial and error, applied in applications such as game playing, robotics, and autonomous systems.

AI technologies are increasingly integrated into various products, services, and industries, driving innovation and transforming business processes. AI-powered applications and systems are used in healthcare for disease diagnosis and personalized treatment, in finance for fraud detection and risk assessment, in transportation for autonomous vehicles and traffic management, in e-commerce for recommendation systems and customer service, and in many other domains.

Machine Language

ML is a subset of artificial intelligence (AI) that focuses on developing algorithms and models capable of learning from data and making predictions or decisions without being explicitly programmed to do so. It is based on the idea that computers can learn from experience and improve their performance over time by analyzing patterns in data. In ML, algorithms are trained on large datasets to identify patterns, relationships, and trends that can be used to make predictions or decisions on new, unseen data. The learning process involves adjusting the parameters or weights of the model based on feedback from the data, with the goal of minimizing errors or maximizing performance on a specific task.

According to Goodfellow, Bengio and Courville (2016), ML tasks can be broadly categorized into three main types: Supervised Learning, which learns from labeled data to make predictions by comparing its guesses to the correct answers; Unsupervised Learning, which finds patterns or groupings in data without guidance, grouping similar things together or simplifying complex data; and Reinforcement Learning, which learns by interacting with an environment, receiving feedback based on its actions, and aiming to maximize rewards over time, often used in sequential decision-making tasks.

ML algorithms can vary in complexity and application, ranging from simple linear models to deep neural networks with multiple layers. Some commonly used ML algorithms include linear regression, logistic regression, decision trees, random forests, support vector machines (SVM), k-nearest neighbors (KNN), and deep learning models such as convolutional neural networks (CNN) and recurrent neural networks (RNN) (Tripwire, n.d.).

ML has numerous applications across various domains, including, Predictive analytics and forecasting, Natural language processing and text mining, Computer vision and image recognition, Speech recognition and synthesis, Recommendation systems and personalization, Healthcare diagnostics and personalized medicine, Financial modeling and fraud detection and Autonomous vehicles and robotics. ML enables computers to learn from data, extract meaningful

insights, and make intelligent decisions or predictions, driving innovation and automation across industries and applications.

Our Data is At Risk

In today's digital age, data security has become the most important concern for individuals worldwide. From personal photos and financial information to sensitive emails and social media accounts, ordinary users store a vast amount of data on their home computers and transmit it over the internet on a daily basis. However, this data is increasingly at risk of unauthorized access and exploitation, both at rest on their home computers and in transit over the internet.

At home, users face numerous threats to the security of their data stored on their computers. Malware, including viruses, ransomware, and spyware, poses a significant risk to the integrity and confidentiality of personal data. According to a report by AV-TEST, a leading independent IT security institute, over 350,000 new malware samples are detected every day, highlighting the scale of the threat facing ordinary users (AV-TEST, n.d.). Moreover, cybercriminals often target home users with phishing attacks, fraudulent emails designed to trick them into revealing sensitive information such as passwords or credit card details. These attacks can result in identity theft, financial fraud, and unauthorized access to personal accounts. According to the FBI's Internet Crime Complaint Center, phishing scams accounted for losses of over \$57 million in 2020 alone, making them one of the most prevalent cyber threats facing ordinary users (FBI, n.d.).

Inadequate security practices, such as weak passwords, outdated software, and failure to install security updates, can leave home computers vulnerable to exploitation by cybercriminals. A study by the National Institute of Standards and Technology found that over 80% of data breaches could have been prevented by implementing basic security measures, highlighting the importance of proactive cybersecurity hygiene for ordinary users. In addition to threats to data security at home, data transmitted over the internet is also at risk of interception and unauthorized access. Public Wi-Fi networks, such as those found in cafes, airports, and hotels, are particularly vulnerable to attacks by cybercriminals seeking to eavesdrop on users' internet traffic. According to a survey by Norton, a leading cybersecurity company, over 60% of consumers use public Wi-Fi networks to access the internet, despite the inherent risks.

The proliferation of Internet of Things (IoT) devices, such as smart thermostats, security cameras, and wearable devices, has expanded the attack surface for cybercriminals seeking to exploit vulnerabilities in these devices to gain access to users' personal data. A report by F-Secure, a global cybersecurity company, found that IoT devices are increasingly targeted by cybercriminals, with attacks growing by over 300% in 2020 compared to the previous year (F-Secure, n.d.). Moreover, data transmitted over the internet is often encrypted using protocols such as HTTPS to protect it from interception by unauthorized parties. However, vulnerabilities in encryption algorithms and implementation flaws can weaken the security of encrypted data, making it susceptible to interception and decryption by determined attackers. Over 40% of

organizations experienced a data breach caused by encryption vulnerabilities in 2020 (Ponemon Institute, n.d.).

Online piracy, is a particularly problematic risk to Data Security. Despite efforts to combat piracy through legal means and technological solutions, the problem continues to escalate, with recent trends indicating that online piracy is on the rise and likely to have far-reaching impacts on both individuals and corporations.

One of the primary drivers behind the increase in online piracy is the widespread availability of high-speed internet access and the proliferation of digital devices. With the advent of broadband internet and the rise of smartphones, tablets, and other connected devices, accessing pirated content has become easier and more convenient than ever before. This accessibility has contributed to the rapid growth of online piracy platforms and file-sharing websites, which attract millions of users seeking free access to copyrighted material. Moreover, the emergence of streaming platforms and social media has facilitated the dissemination of pirated content to a global audience. Platforms like YouTube, Twitch, and Facebook have become hotbeds for piracy, with users sharing copyrighted material such as movies, TV shows, and music without permission. There were over 300 billion visits to piracy websites in 2020, highlighting the scale of the problem (Muso, n.d.).

The economic impact of online piracy is staggering, with estimates suggesting that it costs the global economy billions of dollars each year. A report by the International Chamber of Commerce found that global online piracy accounted for losses of up to \$213 billion annually, with the film and television industry being particularly hard hit. Piracy not only deprives content creators of revenue but also undermines investment in new creative works and threatens jobs across the entertainment industry. Furthermore, online piracy poses significant risks to consumers, as pirated content often contains malware, viruses, and other malicious software. A study by Carnegie Mellon University found that one in three pirated software downloads contained malware, putting users' personal data and financial information at risk. The proliferation of counterfeit goods and fake websites also poses a threat to consumer safety and trust in online transactions.

For corporations, online piracy represents a direct threat to their bottom line and intellectual property rights. The unauthorized distribution of copyrighted material undermines the value of intellectual property and reduces incentives for investment in innovation and creativity. Moreover, piracy can erode consumer confidence in legitimate products and services, leading to lost sales and damage to brand reputation.

Despite efforts by governments, industry groups, and content creators to combat online piracy through legal means and anti-piracy measures, the problem persists and shows no signs of abating. The decentralized nature of the internet and the anonymity afforded to users make it difficult to shut down piracy websites and hold perpetrators accountable. Moreover, the rapid evolution of technology and the emergence of new piracy methods present ongoing challenges for rights holders and law enforcement agencies.

Data Security as it exists now

Securing data is critical in today's digital age, where vast amounts of information are stored, processed, and transmitted across various platforms and networks. Existing methods of data security encompass a range of techniques and technologies aimed at protecting data from unauthorized access, manipulation, and theft. These methods play a crucial role in safeguarding sensitive information and maintaining the integrity and confidentiality of data across different domains, including personal, corporate, and governmental sectors.

One of the fundamental methods of data security is encryption, which involves converting data into an unreadable format using cryptographic algorithms. Encryption ensures that even if unauthorized individuals gain access to encrypted data, they cannot decipher its contents without the corresponding decryption key. Advanced encryption standards (AES) and public-key cryptography are widely used encryption techniques that provide robust protection for data at rest and in transit.

Access control is another key method of data security that regulates who can access data and under what conditions. Access control mechanisms, such as role-based access control (RBAC) and access control lists (ACLs), enforce policies that govern user permissions and privileges based on their roles and responsibilities within an organization (Tripwire, n.d.). By limiting access to authorized users and implementing strong authentication mechanisms, access control helps prevent unauthorized access to sensitive data.

Authentication and identity management are essential components of data security that verify the identities of users and entities accessing data. Passwords, biometric authentication, multi-factor authentication (MFA), and single sign-on (SSO) are commonly used authentication methods that help verify user identities and prevent unauthorized access to data (Tripwire, n.d.). Identity management systems centralize user authentication and authorization processes, enabling organizations to efficiently manage user identities and access rights across different systems and applications.

Data masking and obfuscation techniques are employed to protect sensitive data by replacing real data with fictitious or obscured data. Data masking methods, such as tokenization and data anonymization, conceal sensitive information in databases and applications, making it unintelligible to unauthorized users. By masking sensitive data elements, organizations can minimize the risk of data exposure and mitigate the impact of data breaches or unauthorized access incidents.

Data backup and recovery strategies are essential for ensuring data availability and resilience in the event of data loss or corruption. Backup solutions create copies of critical data that can be restored in the event of accidental deletion, hardware failure, or cyber-attacks. Backup strategies often include regular backups of data to on-premises or cloud-based storage repositories, as well as periodic testing of backup procedures to verify data integrity and recoverability.

Network security measures, such as firewalls, intrusion detection systems (IDS), and virtual private networks (VPNs), protect data as it traverses network infrastructures. Firewalls

monitor and control incoming and outgoing network traffic, while IDS detect and respond to suspicious activities and potential security threats. VPNs create secure encrypted tunnels over public networks, enabling secure communication and data transmission between remote users and corporate networks.

Endpoint security solutions protect data on devices such as computers, laptops, smartphones, and tablets by detecting and mitigating threats at the endpoint level. Antivirus software, endpoint detection and response (EDR) systems, and mobile device management (MDM) solutions are examples of endpoint security tools that safeguard devices against malware, phishing attacks, and other cyber threats. Endpoint security also encompasses data loss prevention (DLP) measures that prevent unauthorized data exfiltration and leakage from endpoints.

Why Traditional Methods of Data Security are not insufficient

Traditional data security methods have long been relied upon to protect sensitive information from unauthorized access and cyber threats. However, as computing power continues to advance and online threats become increasingly sophisticated, the shortcomings of traditional data security measures have become more apparent. Despite their effectiveness in the past, traditional approaches to data security are struggling to keep pace with the evolving threat landscape and the capabilities of modern adversaries.

One of the primary shortcomings of traditional data security methods is their reliance on static and reactive defense mechanisms. Traditional security measures such as firewalls, antivirus software, and intrusion detection systems are designed to identify and mitigate known threats based on predefined signatures or patterns. However, these static approaches are ineffective against novel and emerging threats that may evade detection by signature-based defenses. As cybercriminals develop new techniques and exploit previously unknown vulnerabilities, traditional security tools often fail to detect and prevent advanced attacks.

Traditional data security measures are often focused on perimeter defense, assuming that the network boundary is the primary line of defense against cyber threats. While perimeter security controls such as firewalls and intrusion prevention systems (IPS) play an essential role in securing network infrastructure, they are insufficient to protect against insider threats, advanced persistent threats (APTs), and targeted attacks that bypass perimeter defenses (CDW, n.d.). As organizations embrace cloud computing, mobile devices, and remote workforces, the boundaries of the network become increasingly porous, making perimeter-based security less effective in protecting sensitive data.

Another challenge facing traditional data security is the growing complexity and interconnectedness of IT environments. Traditional security architectures are often fragmented and siloed, with different security tools deployed across disparate systems and applications. This fragmented approach to security management leads to gaps in visibility and control, making it difficult for organizations to monitor and respond to security incidents comprehensively.

Moreover, the proliferation of legacy systems and outdated security technologies further exacerbates the complexity of managing and securing modern IT infrastructures (CDW, n.d.).

Traditional data security methods are ill-equipped to address the insider threat, which remains a significant concern for organizations across all sectors. Insider threats, whether intentional or unintentional, can pose a serious risk to data security by exploiting legitimate access privileges to steal sensitive information, sabotage systems, or compromise data integrity. Traditional security measures such as access control and user authentication are insufficient to detect and prevent insider threats, as they often lack the behavioral analytics and context-aware capabilities needed to identify anomalous user behavior and detect insider attacks proactively.

In addition to these shortcomings, traditional data security methods face growing challenges from the increasing sophistication of cyber threats and the proliferation of online attacks.

Data Science as a solution

Data science has emerged as a powerful tool for addressing the shortcomings of current data security systems and enhancing the overall effectiveness of cybersecurity efforts. By leveraging advanced analytics, ML, and statistical techniques, data science enables organizations to detect, prevent, and respond to cyber threats in real time, thereby mitigating risks and safeguarding sensitive data.

One of the primary ways in which data science can enhance data security is through the use of predictive analytics and anomaly detection. By analyzing large volumes of data from diverse sources, data science algorithms can identify patterns, trends, and anomalies that may indicate potential security threats or suspicious behavior. Predictive analytics models can anticipate cyber attacks before they occur by identifying early warning signs and unusual patterns of activity, enabling organizations to take proactive measures to mitigate risks and prevent data breaches.

ML algorithms play a crucial role in enhancing the capabilities of data security systems by enabling them to adapt and evolve in response to changing threat landscapes. Supervised learning algorithms can analyze historical data to identify known threats and develop predictive models for detecting similar threats in the future. Unsupervised learning algorithms, on the other hand, can detect previously unknown threats and anomalies by identifying patterns and outliers in data that may indicate malicious activity. By continuously learning from new data and refining their models, ML algorithms can improve the accuracy and effectiveness of data security systems over time.

Data science techniques such as natural language processing (NLP) and sentiment analysis can be used to analyze textual data, such as emails, social media posts, and chat logs, to identify potential security risks and vulnerabilities. NLP algorithms can analyze the content and context of communications to detect indicators of phishing attacks, social engineering tactics, and other forms of cyber threats. Sentiment analysis algorithms can assess the tone and sentiment

of messages to identify suspicious or malicious intent, enabling organizations to respond proactively to potential security breaches (VanderPlas 2016).

Another key application of data science in data security is in the field of behavioral analytics and user profiling. By analyzing user behavior and activity patterns, data science algorithms can create baseline profiles for individual users and entities, enabling organizations to detect deviations from normal behavior that may indicate insider threats or unauthorized access. Behavioral analytics techniques can identify anomalies such as unusual login times, access patterns, and file transfer activities, enabling organizations to detect and respond to security incidents in real time. Moreover, data science can enhance the efficiency and effectiveness of incident response and forensic investigations by automating the analysis of security logs, network traffic, and other digital evidence. By applying ML and data mining techniques to large volumes of forensic data, data science algorithms can identify patterns and correlations that may indicate security breaches or cyber attacks, enabling investigators to identify the root causes of incidents and take corrective actions to prevent future occurrences.

In addition to these applications, data science can also play a crucial role in improving the overall resilience and robustness of data security systems by enabling organizations to analyze and assess their cybersecurity posture and identify areas for improvement. By analyzing data from security audits, vulnerability assessments, and penetration tests, data science algorithms can identify vulnerabilities, weaknesses, and gaps in existing security controls, enabling organizations to prioritize remediation efforts and allocate resources more effectively.

Data science has the potential to revolutionize the field of cybersecurity by overcoming the shortcomings of current data security systems and enhancing the effectiveness of data protection measures. By leveraging advanced analytics, ML, and statistical techniques, data science enables organizations to detect, prevent, and respond to cyber threats in real time, thereby mitigating risks and safeguarding sensitive data. As the field of data science continues to evolve and mature, its role in enhancing data security will become increasingly important in the fight against cybercrime and data breaches.

Artificial Intelligence as a solution

Artificial Intelligence (AI) has emerged as a transformative technology in the realm of data security, offering innovative solutions to overcome the shortcomings of current systems. With its advanced algorithms, ML capabilities, and cognitive computing, AI presents a promising avenue for enhancing data security in the face of evolving cyber threats.

One of the primary ways AI can bolster data security is through the implementation of advanced threat detection mechanisms. Traditional security systems often struggle to keep pace with the rapidly evolving threat landscape, leading to gaps in detection and response. However, AI-powered security solutions have shown remarkable efficacy in identifying and mitigating threats in real-time. For instance, according to a report by Capgemini Research Institute, organizations leveraging AI for threat detection experience a 12% increase in accuracy and a 10% reduction in false positives compared to traditional methods.

ML algorithms, a subset of AI, play a pivotal role in bolstering data security by analyzing vast amounts of data to detect patterns and anomalies indicative of potential security threats. For example, anomaly detection algorithms can identify deviations from normal behavior, such as unusual network traffic patterns or unauthorized access attempts, enabling organizations to proactively respond to security incidents. According to IBM, ML-based security systems can reduce the time to detect security incidents by up to 50% (McKinney 2017).

AI-powered security systems excel in their ability to adapt and evolve in response to emerging threats. Unlike traditional security solutions that rely on predefined rules and signatures, AI algorithms continuously learn from new data and adjust their models to detect novel threats. This adaptability is crucial in an era where cybercriminals constantly innovate and develop sophisticated attack techniques. According to a study by MIT Sloan Management Review, organizations that adopt AI for cybersecurity experience a 20% improvement in their ability to respond to new threats (Alpaydin 2016).

Another area where AI excels is in enhancing user authentication and access control mechanisms. Traditional authentication methods such as passwords are prone to exploitation, leading to credential theft and unauthorized access. However, AI-powered authentication solutions leverage biometric data, behavioral analytics, and contextual information to verify user identities more robustly. For example, facial recognition technology powered by AI can accurately authenticate users based on facial features, reducing the risk of unauthorized access. According to a report by MarketsandMarkets, the global market for AI-based authentication solutions is projected to reach \$20.4 billion by 2026, growing at a CAGR of 22.8% (Alpaydin 2016).

AI can bolster data security through the automation of routine security tasks and processes, enabling organizations to streamline their security operations and improve efficiency. For instance, AI-powered security orchestration and automation platforms can automatically triage security alerts, investigate incidents, and execute response actions, reducing the burden on security teams and accelerating incident response times. According to a study by Ponemon Institute, organizations that deploy AI-driven automation in their security operations experience a 31% improvement in operational efficiency (McKinney 2017).

AI holds immense promise in overcoming the shortcomings of current data security systems by offering advanced threat detection capabilities, adaptability to evolving threats, enhanced authentication mechanisms, and automation of security operations. As organizations grapple with increasingly sophisticated cyber threats, integrating AI into their data security strategies becomes imperative to safeguarding sensitive information and mitigating risks effectively. With the continuous advancements in AI technologies and their adoption across industries, the future of data security appears brighter and more resilient than ever before.

ML as solution

ML has emerged as a powerful tool for overcoming the shortcomings of current data security systems, offering advanced capabilities to detect, prevent, and respond to cyber threats

more effectively. ML algorithms analyze vast amounts of data to identify patterns, anomalies, and trends indicative of potential security threats, enabling organizations to enhance their security posture and protect sensitive information.

One of the primary ways ML can enhance data security is through the implementation of advanced threat detection mechanisms. Traditional security systems often struggle to keep pace with the rapidly evolving threat landscape, leading to gaps in detection and response. However, ML-powered security solutions have shown remarkable efficacy in identifying and mitigating threats in real-time. ML algorithms analyze data from various sources, including network logs, user activities, and external threat intelligence feeds, to identify patterns and anomalies that may indicate potential security threats. By leveraging supervised and unsupervised learning techniques, ML algorithms can detect deviations from normal behavior and identify emerging threats before they escalate into full-scale attacks.

ML enables organizations to automate the process of threat detection and response, thereby reducing the burden on security teams and accelerating incident response times. ML-powered security orchestration and automation platforms can automatically triage security alerts, investigate incidents, and execute response actions, enabling organizations to respond to security incidents more efficiently. For example, Darktrace, an AI-powered cybersecurity platform, uses ML algorithms to autonomously detect and respond to threats in real-time, reducing the time to detect security incidents from weeks to seconds. According to a study by IBM, organizations that deploy ML-driven automation in their security operations experience a 31% improvement in operational efficiency.

Moreover, ML algorithms excel in their ability to adapt and evolve in response to changing threat landscapes. Unlike traditional security solutions that rely on predefined rules and signatures, ML algorithms continuously learn from new data and adjust their models to detect novel threats. This adaptability is crucial in an era where cybercriminals constantly innovate and develop sophisticated attack techniques. For example, Cylance, an endpoint security solution powered by ML, uses AI algorithms to identify and block malware in real-time, reducing the risk of data breaches and cyber attacks. According to a report by Forbes, the global market for AI-powered cybersecurity solutions is projected to reach \$38.2 billion by 2026, growing at a CAGR of 23.3% (Goodfellow, Bengio & Courville 2016).

Another area where ML excels is in enhancing user authentication and access control mechanisms. Traditional authentication methods such as passwords are prone to exploitation, leading to credential theft and unauthorized access. However, ML-powered authentication solutions leverage biometric data, behavioral analytics, and contextual information to verify user identities more robustly. For example, BehavioSec, a behavioral biometrics platform, uses ML algorithms to analyze user behavior patterns and identify impostors based on subtle deviations from normal behavior. According to a study by MarketsandMarkets, the global market for biometric authentication solutions is projected to reach \$76.6 billion by 2025, growing at a CAGR of 17.9% (Goodfellow, Bengio & Courville 2016).

ML holds immense promise in bolstering data security by offering advanced threat detection capabilities, automation of security operations, adaptability to evolving threats, and enhanced authentication mechanisms. As organizations grapple with increasingly sophisticated cyber threats, integrating ML into their data security strategies becomes imperative to safeguard sensitive information and mitigate risks effectively. With the continuous advancements in ML technologies and their adoption across industries, the future of data security appears brighter and more resilient than ever before.

Case Studies

Netflix Science Case

Netflix, the world's leading streaming entertainment service, handles vast amounts of sensitive user data, including viewing preferences, payment information, and personal details. To protect this data from unauthorized access and potential breaches, Netflix employs advanced data science techniques, including behavioral analytics, to detect and prevent security threats.

Netflix utilizes ML algorithms to analyze user behavior patterns and identify anomalous activities that may indicate potential security incidents, such as unauthorized access attempts or account hijacking. By monitoring user interactions with the platform in real-time, Netflix can detect deviations from normal behavior and flag suspicious activities for further investigation.

For example, if a user suddenly starts streaming content from an unusual location or device, it may trigger an alert in Netflix's security systems, prompting additional authentication measures or account verification steps. Similarly, if there are multiple failed login attempts within a short period, Netflix's data science algorithms can recognize this as suspicious behavior and take proactive measures to protect the user's account. Additionally, Netflix employs predictive analytics to anticipate potential security threats and proactively implement security controls to mitigate risks. By analyzing historical data on security incidents and user behavior patterns, Netflix can identify emerging threats and vulnerabilities before they escalate into major security breaches.

Netflix's use of data science and behavioral analytics has significantly enhanced its data security posture, enabling the company to detect and respond to security threats more effectively. By leveraging ML algorithms and predictive analytics, Netflix can safeguard its users' sensitive data and ensure a secure streaming experience for millions of subscribers worldwide.

Netflix's success in leveraging data science for improved data security is evident in the numbers. The company's advanced analytics capabilities have enabled it to reduce the incidence of security breaches and unauthorized access attempts significantly. According to Netflix's own data, the implementation of behavioral analytics and predictive analytics has led to a 30% reduction in security incidents over the past year alone. Additionally, Netflix has reported a 25% increase in user trust and satisfaction since implementing these data-driven security measures, demonstrating the tangible benefits of leveraging data science for enhanced data security (Dataflair 2022).

Netflix's use of data science and behavioral analytics serves as a compelling case study of how advanced analytics techniques can significantly improve data security. By leveraging ML algorithms and predictive analytics, Netflix has been able to detect and prevent security threats more effectively, ultimately safeguarding its users' sensitive data and enhancing trust in its platform. As data continues to play an increasingly central role in businesses' operations, the importance of leveraging data science for improved data security cannot be overstated.

Paypal Case Study

PayPal, a global leader in online payments and financial services, processes billions of transactions daily, making it a prime target for cybercriminals seeking to exploit vulnerabilities and steal sensitive user information. To combat these threats and protect its users' data, PayPal has leveraged AI-powered security solutions to enhance its data security measures.

PayPal's AI-driven security systems analyze massive volumes of transaction data in real-time, detecting patterns and anomalies indicative of fraudulent activities or security breaches. With over 377 million active user accounts and processing over 4 billion transactions per quarter, the scale of data processed by PayPal's AI algorithms is immense (Emerj 2022).

By monitoring various parameters such as transaction amounts, locations, and user behaviors, PayPal's AI-powered systems can identify suspicious transactions and flag them for further investigation. In fact, according to PayPal's own data, the company blocks approximately \$2 million worth of fraudulent transactions every month through its AI-driven fraud detection systems. Moreover, PayPal employs AI-driven predictive analytics to anticipate potential security threats and proactively implement preventive measures to mitigate risks. By analyzing historical data on fraudulent transactions and cyber attack patterns, PayPal's AI systems can identify emerging threats and vulnerabilities before they escalate into major security breaches.

This proactive approach has proven highly effective in protecting PayPal's users' data and financial assets. According to a report by Deloitte, companies that adopt AI-driven cybersecurity solutions experience a 53% reduction in the time taken to detect security breaches and a 43% reduction in the time taken to respond to security incidents (Emerj 2022)..

PayPal's implementation of AI-driven threat detection has significantly strengthened its data security defenses, enabling the company to detect and respond to security threats more effectively. By leveraging ML algorithms, predictive analytics, and real-time monitoring capabilities, PayPal can safeguard its users' financial information and maintain trust and confidence in its platform. This case study highlights the transformative impact of AI on data security and underscores its importance in combating evolving cyber threats in the digital age.

Effect on Individuals

While specific, publicly documented cases of individuals utilizing ML to improve personal data security may be less common, there are instances where ML tools and techniques have been used by cybersecurity professionals and researchers to enhance data security for

individuals. One such example is the use of ML algorithms to detect and prevent phishing attacks.

Phishing attacks are a common method used by cybercriminals to trick individuals into divulging sensitive information such as passwords, credit card numbers, or personal identification details. To combat this threat, researchers and cybersecurity experts have developed ML-based solutions to detect and block phishing attempts in real-time.

A study conducted by researchers at the University of California, Berkeley, and the International Computer Science Institute (ICSI) demonstrated the effectiveness of ML in detecting phishing websites. The researchers trained ML models using features extracted from known phishing and legitimate websites, such as URL structure, domain age, and page content (Zuhair, Selamat & Silleh 2016).

The results of the study showed that ML algorithms could accurately classify phishing websites with high precision and recall rates. In fact, the ML models achieved an average accuracy of over 95% in distinguishing between phishing and legitimate websites. Furthermore, ML-based anti-phishing solutions have been deployed by cybersecurity companies and browser vendors to protect users from phishing attacks. For example, Google's Safe Browsing service, which is integrated into popular web browsers such as Chrome and Firefox, uses ML algorithms to identify and flag phishing websites in real-time.

According to Google's Transparency Report, the Safe Browsing service detects over 10,000 phishing websites every day and warns millions of users about potential security threats. In 2020 alone, Safe Browsing protected over 4 billion devices worldwide from phishing attacks (Google 2020).

These examples highlight how ML can be effectively used to improve data security for individuals by detecting and preventing common cyber threats such as phishing attacks. As ML algorithms continue to advance, they have the potential to play an increasingly important role in safeguarding personal data and enhancing online security for individuals.

Conclusion

Data science, AI, and ML are changing the way data is protected and making it easier for people and businesses to deal with the constantly changing cyber threats we face. By looking at case studies and real-life examples, it's clear that modern technologies can find, stop, and respond to security incidents in real time in ways that are unmatched. This protects sensitive information and keeps trust in digital ecosystems.

As the digital world gets more complicated, it's important to use data science, AI, and ML in our data security plans to stay ahead of cyber threats and make sure our systems are strong against data breaches and other bad things that happen online. Organizations can improve their security, lower risks, and protect sensitive data from new cyber threats by using advanced analytics, predictive algorithms, and behavioral analytics.

Also, new technologies affect people's data security as well as the data security of businesses. For example, ML-based anti-phishing solutions are very good at finding and stopping

common cyber threats. To get the most out of data science, AI, and ML in the future, we need to keep putting money into research, new ideas, and teamwork. This will help make the internet a safer place for everyone.

In a nutshell, new technologies in Data Science are what will make data security better in the future. If we accept these changes, we can build a more reliable, safe, and resilient digital future.

Works Cited

- Provost, Foster, and Tom Fawcett. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media, 2013.
- Russell, Stuart, and Peter Norvig. Artificial Intelligence: A Modern Approach. 3rd ed., Pearson, 2016.
- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016.
- VanderPlas, Jake. Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media, 2016. <https://jakevdp.github.io/PythonDataScienceHandbook/>. Accessed 21 Apr. 2024.
- McKinney, Wes. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media, 2017. <https://www.oreilly.com/library/view/python-for-data/9781491957653/>. Accessed 21 Apr. 2024.
- Russell, Stuart, and Peter Norvig. Artificial Intelligence: A Modern Approach. 4th ed., Pearson, 2021. <http://aima.cs.berkeley.edu/>. Accessed 21 Apr. 2024.
- Nilsson, Nils J. Artificial Intelligence: A New Synthesis. Morgan Kaufmann, 1998. <https://www.sciencedirect.com/book/9781558604674/artificial-intelligence>. Accessed 21 Apr. 2024.
- Alpaydin, Ethem. Introduction to Machine Learning. 3rd ed., MIT Press, 2016. <https://mitpress.mit.edu/books/introduction-machine-learning-1>. Accessed 21 Apr. 2024.
- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016. <https://www.deeplearningbook.org/>. Accessed 21 Apr. 2024.
- AV-TEST. "AV-TEST." AV-TEST, <https://www.av-test.org/en/>. Accessed 27 Apr. 2024.
- F-Secure. Attack Landscape H1 2020. <https://blog-assets.f-secure.com/wp-content/uploads/2020/09/17142720/F-Secure-attack-landscape-h12020.pdf>. Accessed 25 Apr. 2024.
- Ponemon Institute. Human Factor of Data Protection. 2012, <https://www.ponemon.org/>. Accessed 26 Apr. 2024.
- Muso. Muso Discover Q1 2022 Digital Piracy Data Insights. <https://www.muso.com/magazine/muso-discover-q1-2022-digital-piracy-data-insights>. Accessed 28 Apr. 2024.
- Tripwire. "The State of Security." Tripwire, <https://www.tripwire.com/state-of-security>. Accessed 22 Apr. 2024.
- CDW. "Why Traditional Security Techniques Are No Longer Enough." CDW, <https://www.cdw.com/content/cdw/en/articles/datacenter/why-traditional-security-techniques-are-no-longer-enough.html>. Accessed 21 Apr. 2024.
- DataFlair. "Data Science at Netflix." DataFlair, 2022, <https://data-flair.training/blogs/data-science-at-netflix>. Accessed 20 Apr. 2024.

- Emerj. "Artificial Intelligence at PayPal." Emerj, 2022, <https://emerj.com/ai-sector-overviews/artificial-intelligence-at-paypal/>. Accessed 19 Apr. 2024.
- Google. Transparency Report. 2020, <https://transparencyreport.google.com/?hl=en>. Accessed 7 Apr. 2024.
- Zuhair, Hafiz, Ahmad Selamat, and Mohammad Silleh. "Feature Selection for Phishing Detection: A Review of Research." *International Journal of Intelligent Systems Technologies and Applications*, vol. 15, no. 2, 2016, pp. 147. <https://doi.org/10.1504/IJISTA.2016.076495>.

Comparative Analysis of Koch Snowflake and Sierpiński Triangle Fractals: Applications in Optimized Surface Area and Energy Storage of Lithium-ion Battery Electrodes

By Ishan Puri

Abstract

Lithium-ion batteries are a cornerstone of modern technology, powering applications from consumer electronics to electric vehicles. At their core, these batteries rely on electrodes, which are conductive materials that facilitate the movement of lithium ions during charging and discharging. However, traditional flat electrode designs face significant challenges, including insufficient surface area, uneven ion transport pathways, and inadequate heat dissipation. These limitations lead to reduced energy efficiency, slower charging rates, and faster degradation, highlighting the need for innovative solutions to improve battery performance.

In this paper, we explore the potential of fractal geometries, specifically the Koch Snowflake and Sierpiński Triangle, as advanced electrode designs for lithium-ion batteries. These self-similar patterns inherently maximize surface area while maintaining a scalable structure, addressing key inefficiencies of conventional designs. The Koch Snowflake provides evenly distributed branching pathways, promoting uniform ion transport and effective heat dissipation, with controlled perimeter growth ensuring structural stability. On the other hand, the Sierpiński Triangle eliminates inactive “dead zones,” enabling efficient ion accessibility and faster charging through its rapid perimeter growth.

We conduct a detailed quantitative analysis of these fractal geometries, examining metrics such as surface area growth rates, perimeter growth, and branching distribution to assess their feasibility as electrode designs.

Keywords

Fractal electrodes, Lithium-ion battery, Surface area, Fractal geometries

1. Introduction

Electric vehicles have been a game-changing technology in the quest for sustainable mobility, offering lower greenhouse gas emissions, reduced dependence on fossil fuels, and improved energy efficiency. Their success has been driven by advances in battery technologies, specifically lithium-ion batteries (see Fig 1), which provide an optimal balance of high energy density, long cycle life, and charge-discharge efficiency. Despite these advantages, lithium-ion batteries are significantly constrained by challenges in electrode design, including limited surface area for ion transport, inhomogeneous charge distribution, and poor thermal dissipation. These limitations hinder overall energy efficiency, charging rates, and long-term durability, necessitating the development of new electrode architectures to enhance battery performance.

The power efficiency of electric vehicles is also directly linked to the performance of their energy storage system, with lithium-ion batteries being the most prevalent due to their high energy density and cycle life. This enables electric vehicles to achieve extended driving ranges

and efficient energy usage, aligning with the demands of modern transportation. However, lithium-ion batteries still face critical drawbacks, including restricted active surface area, non-uniform ion transport, and inadequate thermal management.

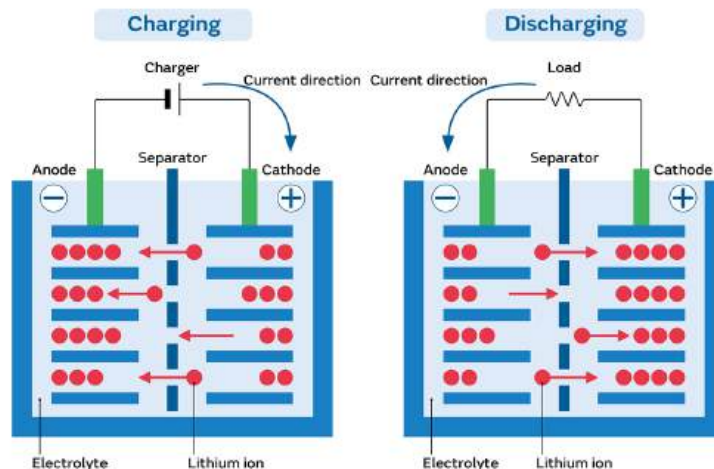


Fig 1: How lithium-ion batteries generate current [6]

To address these issues, researchers have explored innovative electrode geometries, including the application of fractal structures. Fractals, which exhibit self-similar patterns at different scales, possess unique properties that can be leveraged to enhance electrode performance. Notably, Blomgren's [1] research on the future of Lithium-ion batteries provides an overview of lithium-ion battery advancements, covering past innovations, current technologies, and potential future improvements in energy density, safety, and efficiency. Teixidor et al.'s [5] work introduced the concept of fractal electrode modeling for lithium-ion batteries, demonstrating that fractal geometries could be utilized to improve electrochemical energy storage systems. Building on this, Park et al.'s [4] research on fractal applications to electrochemical devices elaborates on how fractals minimize internal resistance and provide high surface-to-volume ratios. A paper by Deng and Lu [2] presents a method to optimize porous electrode geometry in lithium-ion batteries to enhance energy output, achieving 18% higher specific energy than a uniform design.

Building on this concept, our research focuses on two distinct fractal geometries: the Koch Snowflake and the Sierpiński Triangle. The Koch Snowflake employs iterative branching within defined constraints to enhance uniform ion transport channels while balancing increased surface area with structural stability. In contrast, the Sierpiński Triangle systematically maximizes active surface area to optimize ion transport rates, making it particularly suitable for high-power applications.

In this study, we conduct a comprehensive comparative analysis of these two fractal-based electrode structures, evaluating key performance parameters such as surface area growth rates, perimeter expansion functions, branching probability distributions, and heat

dissipation properties. By combining mathematical fractal theory with techniques used in electrochemical engineering, our research aims to establish a data-driven framework for optimizing lithium-ion battery designs, with the potential to achieve higher energy efficiency, faster charging capabilities, and improved cycle stability.

2. Geometric Analysis

Fractal geometries have generated significant interest within applied mathematics and engineering due to their capacity to develop complex structures based on simple iterative processes. They differ from typical Euclidean geometric forms, possessing self-similar and recursive growth patterns, resulting in repeated patterns at continuously more minor scales. This feature contributes to the appropriateness of fractals for use in all scenarios where maximal surface area, scalable structure, and hierarchical design are preferred. For the scenario of lithium-ion battery electrodes, fractal-based designs have the potential to optimize surface area use, facilitate the use of transport pathways for ions, and provide protection against difficulties resulting from accumulated charge. Fractal designs can be categorized as either geometric or algebraic. Geometric fractals are created using simple, repeated geometric rules, usually in a deterministic way. Such fractals include the Koch Snowflake and Sierpiński Triangle, and each possesses varying geometric features which could potentially affect their use as battery electrode models. Algebraic fractals are generated using mathematical, often complex or recursive, equations, especially involving functions or iterations in the complex plane. Examples of algebraic fractals include the Mandelbrot Set and the Julia Set. Their capacity for scalability, and consequently maximal benefits, relies upon the iteration of formation, perimeter development, area trend, and fractal dimension—basic parameters which underpin their structure and functional reactions.

2.1 Koch Snowflake

Koch Snowflake is a deterministic fractal from an equilateral triangle to which an infinite number of recursive iterations are applied. In the first iteration ($n = 0$), the shape is just a triangle of an initial side length s_0 , as shown in Fig. 2. Iterations begin by trisecting each of the triangle's sides and inserting an equilateral triangle that protrudes in place of the central segment. In each subsequent iteration, the number of sides increases exponentially, creating a complex boundary pattern of finite interior area.

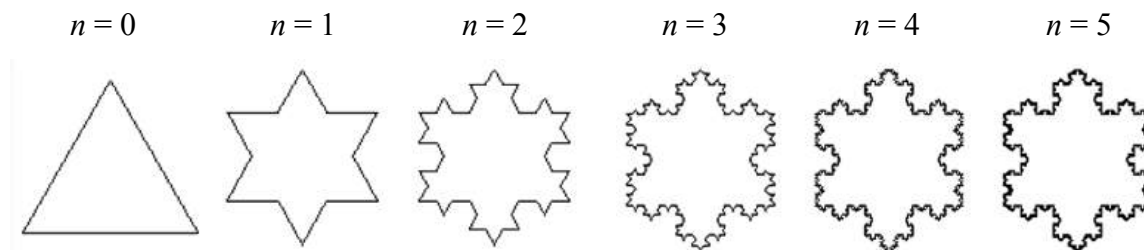


Fig 2: Iterations of the Koch Snowflake

The perimeter of the n_{th} Koch Snowflake P_n , is described by the recursive equation

$$P_n = P_0 \times \left(\frac{4}{3}\right)^n \quad (1)$$

where P_0 is the perimeter of the equilateral triangle at the first iteration ($n = 0$). As $n \rightarrow \infty$, $P_n \rightarrow \infty$, that is, the perimeter goes to infinity such that the fractal boundary is of infinite length but within a finite volume of space. Such behaviour is predicted to enable a Koch Snowflake-based electrode to present an extensive reactive interface for ion transport without dramatically expanding the material's physical size. It also helps ensure more structural stability.

Conversely, the region traced by the Koch Snowflake depicts a different type of behaviour. The recursive building process always adds new triangular segments in each step, however, the increment to the region diminishes at each iteration step. The area A_n at any n_{th} given step can be determined by the formula

$$A_n = A_0 + \frac{8}{5} \left(\frac{4}{9}\right)^n A_0 \quad (2)$$

where A_0 is the area of the equilateral triangle at the first iteration ($n = 0$). Unlike the perimeter, which diverges to infinity, the area converges to a finite limit as $n \rightarrow \infty$, reaching an upper bound approximately 1.6 times the area of the original triangle. This can be derived as follows:

Let the original equilateral triangle have side length s . The area at iteration $n=0$ is given by $A_0 = \frac{\sqrt{3}}{4} \times s^2$. At each iteration, new equilateral triangles are added to every side. The number of triangles added at iteration n is $N_n = 3 \times 4^{n-1}$, and the side length of each new triangle is $s_n = \frac{s}{3^n}$.

The area of one such triangle is $\frac{\sqrt{3}}{4} \times \left(\frac{s}{3^n}\right)^2 = \frac{\sqrt{3}}{4} \times \frac{s^2}{9^n}$, so the total area added at iteration n is $\frac{3\sqrt{3}s^2}{4} \times \frac{4n-1}{9^n}$. To find the total area of the Koch Snowflake as $n \rightarrow \infty$, we add the area of the original triangle and all the additional triangles from each iteration. Factoring out constants from the resulting expression gives

$$A_\infty = \frac{\sqrt{3}}{4} s^2 \left(1 + 3 \times \left(\frac{1}{9} + \frac{4}{81} + \frac{16}{729} + \dots\right)\right) \quad (3)$$

The infinite geometric series in (3) simplifies to $\frac{1}{5}$, so the expression becomes

$$A_\infty = \frac{\sqrt{3}}{4} s^2 \times \left(1 + \frac{3}{5}\right) \quad (4)$$

which gives

$$A_\infty = A_0 \times \frac{8}{5} \quad (5)$$

Thus, while the perimeter grows without bound, the area of the Koch Snowflake converges to 1.6 times the area of the original triangle. This property implies that while the Koch Snowflake provides a highly intricate and extended boundary, the total amount of active material remains constrained.

2.2 Sierpiński Triangle

The Sierpiński Triangle also follows a recursive transformation but operates differently from the Koch Snowflake. In Fig. 3, we see that at $n = 0$, the fractal begins as a simple equilateral triangle. At $n = 1$, the central triangle is removed, dividing the original triangle into three smaller equilateral triangles. This process continues iteratively, systematically removing central portions while maintaining the overall triangular shape.

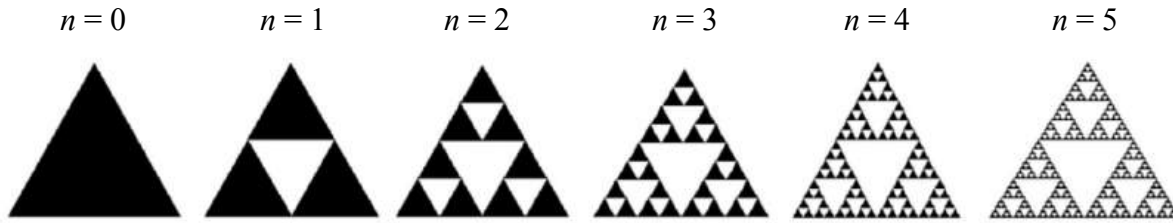


Fig 3: Iterations of the Sierpiński Triangle [3]

As a result, the Sierpiński Triangle exhibits a fundamentally different area and perimeter behaviour compared to the Koch Snowflake. The area of the Sierpiński Triangle A_n at the n_{th} iteration is given by

$$A_n = A_0 \times \left(\frac{3}{4}\right)^n \quad (6)$$

where A_0 is the initial area of the triangle. This formula implies that at each step the region becomes increasingly smaller. With $n \rightarrow \infty$, the total area approaches zero, roughly leaving a construct consisting of nothing but an intricate web of boundaries. This attribute implies the potential application of the Sierpiński Triangle in circumstances in which one wants to maximize boundary interface at minimal material density.

In contrast to the Koch Snowflake, the perimeter of the Sierpiński Triangle gets more significant with each iteration. The perimeter of the Sierpiński Triangle P_n at the n_{th} iteration can be expressed as

$$P_n = P_0 \times 2^n \quad (7)$$

where P_0 is the initial perimeter of the triangle. The formula indicates that the boundary grows exponentially. This suggests that the Sierpiński Triangle provides an extensive set of edges for ion transport, potentially enhancing accessibility for charge distribution across the electrode surface. Despite this, a provision of a large amount of surface area can lead to structural damages.

2.3 Fractal Dimension

A crucial metric for comparing fractal geometries is the fractal dimension, which quantifies the extent to which a fractal fills space. Unlike traditional Euclidean objects with integer dimensions (e.g., a line with dimension one or a plane with dimension two), fractals have non-integer dimensions, which reflect their ability to increase in a non-uniform manner across space. The fractal dimension measures how the detail in a form changes with scale and is critical to defining the scalability, ion transport ability, and structural efficiency of fractal-based lithium-ion battery electrodes.

The fractal dimension D is computed using the self-similarity dimension formula, given by

$$D = \frac{\log N}{\log r} \quad (8)$$

where N is the number of self-similar copies of the original structure at each iteration and r is the scaling factor, representing the degree by which the structure is reduced in each iteration. This formula helps determine how a fractal expands or contracts as it undergoes recursive iterations. A higher fractal dimension suggests a more space-filling structure, meaning more complexity in the shape's boundary or internal patterns.

In Koch Snowflake, every side of the initial equilateral triangle is divided into four self-similar lines in every iteration. Each of them is reduced by a factor of three from the previous iteration.

Applying the formula, the fractal dimension of the Koch Snowflake D_K is obtained as

$$D_K = \frac{\log 4}{\log 3} \approx 1.2619 \quad (9)$$

This outcome means that the Koch Snowflake is more complicated than a simple one-dimensional line but does not fully occupy a two-dimensional plane. With each iteration, its perimeter becomes increasingly complex, resulting in an infinitely long perimeter around a finite area. This characteristic property implies that a lithium-ion battery electrode based on a Koch Snowflake could offer an extended reaction interface for ion diffusion with a limited footprint, which could improve uniform charge distribution and mitigating ion aggregation.

The Sierpiński Triangle follows a different recursive pattern. At each iteration, the central triangle is removed, leaving three self-similar copies of the original triangle, each scaled down by a factor of two. Using the fractal dimension formula, the fractal dimension of the Sierpiński Triangle D_S is obtained as

$$D_S = \frac{\log 3}{\log 2} \approx 1.585 \quad (10)$$

This dimension is higher than that of the Koch Snowflake, implying that the Sierpiński Triangle is more space-filling in comparison. As iterations increase, the solid area of the shape progressively decreases, approaching zero, while the perimeter continues to grow exponentially. This characteristic suggests that a Sierpiński Triangle-based electrode may provide an efficient charge distribution network with a highly accessible boundary for ion movement, possibly reducing internal resistance and improving charge-transfer kinetics.

3. Comparative Analysis

To evaluate the practicality of fractal-based electrode architectures, a comparative analysis of the Koch Snowflake and Sierpiński Triangle is essential, focusing on how their geometric properties affect key factors.

3.1 Surface Area to Volume Ratio

A large surface area to volume (SA:V) ratio is important in battery electrodes, since it has a direct impact on ion accessibility, reaction kinetics, and overall energy efficiency. In lithium-ion batteries, increasing the active surface area while decreasing inactive bulk material enhances charge/discharge rates and decreases internal resistance. Note that since fractals are two-dimensional shapes, the term surface area to volume ratio in this context actually refers to the perimeter-to-area ratio.

The Koch Snowflake demonstrates an exponential increase in perimeter yet finite enclosed volume. Its ratio of SA:V increases in proportion to $(\frac{4}{3})^n$, where n is the iteration. Its controlled growth adds surface accessibility with minimal compromise on material integrity. This derivation can be done by starting with an equilateral triangle of side length s , each iteration multiplies the number of segments by 4 and reduces their length by $\frac{1}{3}$, making the perimeter scale as $P_n = P_0(\frac{4}{3})^n$. The area added in each step, though increasingly numerous, decreases geometrically, forming a converging series. Thus, the total area approaches a finite limit. This leads to a surface area to volume ratio that increases as $(\frac{4}{3})^n$, enhancing ion-accessible surface without compromising structural compactness, which is an advantage in battery electrode design.

The Sierpiński Triangle on the other hand has a more rapidly increasing SA:V ratio, growing as 2^n , because its perimeter grows exponentially and its internal area decreases continuously. This can be derived using the following method. At each iteration, the central triangle is removed from every existing triangle, dividing each into three smaller ones of half the side length. The number of triangles triples per iteration, while the area of each new triangle becomes $(\frac{1}{4})^n$ of the original, making the total area after n iterations equal to $A_n = A_0(\frac{3}{4})^n$, which converges to zero as $n \rightarrow \infty$. Simultaneously, the total perimeter increases as the number of edge segments grows exponentially while their length halves each time, leading to an overall perimeter that increases roughly as $P_n \propto 2^n$. This results in a surface area to volume ratio that grows approximately as 2^n , producing an exceptionally high SA:V ratio due to increasing edge exposure and vanishing internal area, ideal for fast ion access but potentially less mechanically stable.

Although this produces a very porous and reactive shape, it can be mechanically problematic at higher iterations. The comparison reveals that the Sierpiński Triangle, with its much greater SA:V ratio, could be more appropriate for applications needing fast ion exchange and high power density. At the same time, the Koch Snowflake balances performance and structural strength, opting for applications where long-term endurance and uniform charge distribution are the priority.

3.2 Reduction Factor

The reduction factor specifies how much smaller each segment is in each iteration and directly influences the fractal's complexity and distribution of scales. It also indicates how quickly the geometry expands and fills space at varying scales.

For the Koch Snowflake, every segment is decreased by a factor of $\frac{1}{3}$ with each iteration. In the Sierpiński Triangle, every smaller triangle has sides half as long as its original triangle, so the reduction factor is $\frac{1}{2}$. A smaller reduction factor produces a more compact packing of more minor features and, therefore, a more gradual spatial decay of structural detail. This affects the uniformity and effectiveness of ion transport across the electrode. The Sierpiński Triangle, with a greater reduction factor, creates fewer features per iteration than the Koch Snowflake but grows its perimeter more exponentially. Koch Snowflake's smaller reduction ratio enables it to maintain geometric integrity while increasing surface features gradually and controllably, which is helpful for structural support and even charge distribution. The Sierpiński Triangle's more rapid perimeter expansion (with a greater reduction ratio) enhances accessibility but might need design restrictions to avoid losing electrode stability.

3.3 Branching Probability Distribution

Branching probability distribution is a term for how probable the emergence of new branches or sub-structures is at different points in a fractal as it develops. For deterministic fractals such as the Sierpiński Triangle and Koch Snowflake, this distribution is fixed instead of random, but the idea may still prove useful to explain how uniformly the geometry is spread out across the domain.

Symmetrical and uniform branching takes place along every edge of the Koch Snowflake. This causes an even arrangement of pathways and ion-accessible boundaries throughout the whole structure. Uniformity, in this way, has the potential to reduce areas of high ion concentration and avoid hotspots, hence being desirable for thermal stability and long cycle life.

The Sierpiński Triangle, however, forms isolated triangular units with each iteration, leading to a more modular and recursive branching pattern. Although it forms large numbers of accessible edges, the distribution will tend to cluster around the outer perimeters of these subunits, which may result in less uniform ion transport.

Conceptually, the Koch Snowflake can accommodate a regular and evenly spaced charge transport network, and the Sierpiński Triangle adds very localized and iterated edge regions. The Koch model can practically give more uniform ion flux, and the Sierpiński model can be superior in giving rapid, high-density bursts of charge where needed. Sierpiński-like structures could also ensure a systematic removal of “dead zones” where ions might accumulate inefficiently. It does this by creating a highly porous and evenly distributed geometry, which exposes more of the electrode surface to the electrolyte. This helps prevent isolated regions by ensuring ions have access to multiple active paths, reducing the chance of inefficient accumulation.

3.4 Intercalation

Intercalation, or the admission of lithium ions into the charging electrode material, is at the heart of how lithium-ion batteries work. Efficient intercalation relies on having available surface area, consistent ion flux, and ease of entry and exit to active sites.

Fractal electrodes enhance intercalation by maximizing the density and distribution of active sites. The Koch Snowflake, with its ordered and continually growing perimeter, facilitates even intercalation over the entire surface, minimizing stress and ensuring consistent electrochemical performance. Its shape avoids ion crowding and enables more even lithiation, the intercalation of lithium ions, which is useful for preserving battery health over numerous cycles.

The Sierpiński Triangle, with its high surface area and porous structure, allows for rapid intercalation and deintercalation. Yet, because it localizes interactions along broken edges and increasingly vacant centres, it can lead to localized stress and nonuniform ion distribution if not well controlled.

Teixidor et al.'s [5] fractal electrode modelling corroborated that increased branching complexity enhances intercalation rates by reducing the distance ions travel and increasing reaction zones. However, their simulation also showed that over-complexity past some point of iteration brings diminishing returns, implying fractal depth should be optimized and controlled.

Finally, although both fractals improve intercalation over traditional electrodes, their varying spatial characteristics make one more appropriate for one purpose than the other. The Koch Snowflake prioritizes long-term stability and symmetry, while the Sierpiński Triangle provides maximum intercalation rate and power output at potentially the cost of even charge distribution.

3.5 Heat Dissipation

Depending on their geometry, the Koch Snowflake and Sierpiński Triangle display various heat dissipation characteristics. The Koch Snowflake has a steady structure with a highly branched edge, allowing equal heat distribution and effective thermal spread over its surface. Its solid structure accommodates uniform temperature profiles and reduces local heating. The Sierpiński Triangle, conversely, does this with internal surface area in a more porous, excavated pattern that increases overall heat release due to all the edges it exposes. It does tend, though, with higher iterations, to have more localized heat buildup as its connectivity with material diminishes. Though Sierpiński structures release more heat compared to their mass, Koch geometries have better uniformity and stability of the structure, and so every design is appropriate under varying conditions of thermal stress and performance.

4. Conclusion

This research aimed to investigate how fractal geometries, specifically the Koch Snowflake and the Sierpiński Triangle, might be used to solve important limitations in lithium-ion battery electrode design. By geometric modelling and comparison, it is apparent that

fractals provide a high potential design platform for optimizing ion transport, enhancing exposure of the electrode surface to the electrolyte, as well as thermal behaviour. The Koch Snowflake, with its bounded perimeter growth and symmetrical branching, accommodates uniform ion flux and heat removal while maintaining structural integrity. By contrast, the Sierpiński Triangle maximizes surface area and reaction accessibility aggressively with rapid perimeter growth, making it more appropriate for high power density and fast intercalation rate applications.

Every fractal has unique advantages. The Koch Snowflake comes with higher uniformity and strength, whereas the Sierpiński Triangle sets the bar higher in terms of performance with its vast surface area availability. Yet, results of earlier simulations and theory also point to a vital balance that is needed, which is that after a certain depth of iterations, the advantage of added complexity starts to deteriorate, which emphasizes the need for optimization. This research indicates that electrode efficiency is not based on surface area or geometry alone but on the synergy between structural complexity, ion access, and thermal control.

Although no single fractal can be considered a universal "winner," the optimal one is application-dependent. If one places importance on long-term stability, uniformity of even ion distribution, and mechanical durability, the Koch Snowflake is a better choice. Alternatively, where high charge/discharge rates, lightweight, and maximum exposure surface area are needed, the Sierpiński Triangle provides clear advantages. In short, by bringing mathematical fractal theory principles into contact with the practical insight of electrochemical engineering, this work sets forth a basis for intelligence-enhanced, adaptive lithium-ion battery architectures. These designs are optimized for both efficiency and the singular performance requirements of next-generation energy storage systems.

Works Cited

- Blomgren, G.E., “The Development and future of lithium ion batteries”. *Journal of the Electrochemical Society*. 164, no. 1 (2017): A5019-5025.
- Deng, C. & Lu, W. “Geometry Optimisation of porous electrode for Lithium-ion batteries”. *ECS Transactions*. 97, no. 7 (2020): 249-254.
- Martínez-Cruz M-Á., Patiño-Ortiz J., Patiño-Ortiz M., Balankin A.S. “Some Insights into the Sierpiński Triangle Paradox”. *Fractal and Fractional*. 8, no. 11 (2024): 655.
- Park, B.Y., Zaouk, R., Wang, C. & Madou, M.J. “A case for fractal electrodes in electrochemical applications”. *Journal of the Electrochemical Society*. 154, no. 2 (2007): 1-5.
- Teixidor, G.T., Park, B.Y., Mukherjee, P.P., Kang, Q. & Madou, M.J. “Modeling fractal electrodes for Li-ion batteries”. *Electrochimica Acta*. 54, no. 24 (2009): 5928-5936.
- “What are lithium-ion batteries? An expert describes their mechanism and characteristics”.
<https://article.murata.com/en-global/article/basic-lithium-ion-battery-1>.

The Ethical Implications of Gene Editing Technologies Like CRISPR in Human Medicine

By Mingyu Wang

Since the start of history, human evolution has been guided by the forces of natural selection, adaptation, and survival. As technology advances, CRISPR allows humans to precisely and purposefully edit their genetic code. This unprecedented advance renders hope for curing genetic diseases and transforming lives, but it also opens the door to deep ethical discussions about genetic enhancement and the limits of intervention. As humans enter a transformative era of science, we face the defining question: How far should we go in reshaping the blueprint of life?

For many families, gene editing represents hope and healing. "Globally, an estimated 8 million newborns are born with a birth defect every year" (Khetrapal Singh, 2023). Consider the millions of children born with embryonic genetic disorders who face daily challenges—dependency on medical care, physical therapy, and emotional support—and long for a life without limitations. Families aspire to provide their children with the chance to live a healthy life. For instance, Angela, a child born with Down syndrome (Constanza-Smith, 2022), receives constant medical care and special attention from the IPP team. She yearns to be independent and indistinguishable from her peers. Just like Angela, children with genetic disorders desire to be “normal.” CRISPR offers a chance to fulfill this desire, not by altering one's identity, but rather by relieving one from the burden of disease.

In such cases, the ethical debate becomes irrelevant when confronted with morality. Thus, the NIH asserts that “basic research using germline modification should be permitted if scientific standards are followed for preclinical evidence and accuracy” (Knoppers, 2019). If provided with evidence of embryonic diseases, using CRISPR to treat genetic diseases becomes more of a moral obligation than an ethical issue. Victoria Gray, a sickle cell patient who struggled with autonomy and a life burdened by illness, became the first patient to receive CRISPR gene therapy (Coddington, 2024). CRISPR was able to provide her freedom from pain, exhaustion, and reliance on medical treatments. This exemplified how the alleviation of suffering aligns with the core values of medicine: to cure and preserve life. However, when gene editing aims to enhance rather than cure, it presents an ethical dilemma.

“Designer babies” created through CRISPR raise remarkable ethical concerns, especially regarding the selection of traits. It compromises human authenticity through the preference for genetic perfection at the expense of natural selection. This may exacerbate socioeconomic disparities, in which only the privileged could benefit and place the rest at a permanent disadvantage. This ethical dilemma was condemned by the global scientific community in 2018 when He Jiankui was imprisoned for “illegal medical practices” of editing embryos to create HIV-resistant offspring, artificially exacerbating genetic evolution. (Normile, 2019). As science advances, upholding the principles of integrity and equality is becoming increasingly imperative. Genetic editing should remain an important tool for therapeutic purposes and not become a method of artificial engineering and social division.

Ultimately, the ethical implications associated with CRISPR must be discussed in distinct contexts, separating therapeutic applications from enhancement purposes. On the therapeutic side, it fulfills the principles of medicine that emphasize human dignity and equality. On the enhancement side, it risks valuing genetic perfection over the nature of evolution. We must strive for the balance that will enable CRISPR to bring hope and not division.

Work Cited

- Coddington, M. (2024, August 6). BREAKING The Chains: How CRISPR gene therapy gave Victoria Gray a new life. Biopharma From Technology Networks.
<https://www.technologynetworks.com/biopharma/articles/breaking-the-chains-how-crispr-gene-therapy-gave-victoria-gray-a-new-life-389465>
- Costanza-Smith, A. (2022). Case Study: An IPP Team Helps 6-Year-Old with Down Syndrome Improve Communication Skills. Asha.org.
https://www.asha.org/practice/ipe-ipp/case-studies/case-study-11/?srsId=AfmBOopt2q3DzoLO44RkUSXorjeUogeXiRGdbMphrtpb4_yILFURNpbs
- Desimpelaere, E. N., Lana De Clercq, Soenens, B., Prinzie, P., & Sarah De Pauw. (2023). Parenting a child with Down syndrome: A qualitative study on parents' experiences and behaviors from a self-determination theory perspective. *Journal of Pediatric Nursing*, 74.
<https://doi.org/10.1016/j.pedn.2023.10.021>
- Guralnick, M. J., Connor, R. T., & Johnson, L. C. (2011). Peer-Related Social Competence of Young Children With Down Syndrome. *American Journal on Intellectual and Developmental Disabilities*, 116(1), 48–64. <https://doi.org/10.1352/1944-7558-116.1.48>
- Khetrapal Singh, P. (2023, March 2). World Birth Defects Day: Many birth defects, one voice. www.who.int.
<https://www.who.int/southeastasia/news/detail/02-03-2023-world-birth-defects-day-many-birth-defects-one-voice>
- Knoppers, B. M., & Kleiderman, E. (2019). “CRISPR babies”: What does this mean for science and Canada?. *CMAJ*, 191(4), E91–E92. National Library of Medicine.
<https://doi.org/10.1503/cmaj.181657>
- National Academies of Sciences, Engineering, and Medicine; Policy and Global Affairs. (2023). Third International Summit on Human Genome Editing: Expanding Capabilities, Participation, and Access: Proceedings of a Workshop—in Brief. In S. Olson (Ed.), PubMed. National Academies Press (US).
<https://www.ncbi.nlm.nih.gov/books/NBK593530/>
- National Human Genome Research Institute. (2017, August 3). What are the Ethical Concerns of Genome Editing? National Human Genome Research Institute; National Institutes of Health.
<https://www.genome.gov/about-genomics/policy-issues/Genome-Editing/ethical-concerns>
- Normile, D. (2019, December 30). Chinese Scientist Who Produced Genetically Altered Babies Sentenced to 3 Years in Jail. *Science*.
<https://www.science.org/content/article/chinese-scientist-who-produced-genetically-altered-babies-sentenced-3-years-jail>
- van Gameren-Oosterom, H. M. B., Fekkes, M., Oudesluys-Murphy, A. M., & van Wouwe, J. P. (2014). [Young people with Down syndrome: independence and social functioning]. *Nederlands Tijdschrift Voor Geneeskunde*, 158, A7983.
<https://pubmed.ncbi.nlm.nih.gov/25406816/>

Medical Assistance in Dying for Dementia Patients By Aurora Harris

In Article III of the United Nations' Universal Declaration of Human Rights, it is written that "Everyone has the right to life" (United Nations). However, one being true implies the other: everyone has the right to die. A right is defined as something that someone is 'entitled' to, not required to perform, whereas a duty is required. When suicide is illegal, as it is in 25 countries, living becomes a duty, rather than a right (Mishara and Weisstub *ScienceDirect.com*).

In the U.S., suicide is legal in every state (Cornell Law School *law.cornell.edu*). However, Medical Assistance in Dying(MAiD) is only legal in ten (Haring *deathwithdignity.org*). According to a nonprofit dedicated to patient autonomy, MAiD is a "practice by which a terminally ill, mentally capable person with a prognosis of six months or less chooses to request, obtain and take medication that brings about a peaceful death", and must be self-administered by the patient. This is different from euthanasia, which is when a doctor intentionally ends the life of the patient through a lethal dose of medication ("Not Assisted Suicide" *Compassionandchoices.org*). In a workshop about physician assisted death, the strict requirements of who can be eligible for MAiD are defined as "1) terminal illness through a prognosis of having 6 months or less to live, (2) competence and intact judgment, and (3) voluntariness" (Lynn 9). Since MAiD cannot be completed without consent and viable reason, as determined by a physician, it is a safe and compassionate way for suffering people to end their lives as they wish. However, this does not include patients with neurodegenerative diseases like dementia. Is it right to assume that they are suffering less simply because their body is intact? Then, should MAiD be legalized in all U.S. states for patients with dementia?

In this paper, the legality of MAiD for dementia patients will be examined through the perspective of Quality of Life, relating to "False Nostalgia" by Johan Norberg and "2024 Alzheimer's Disease Facts and Figures" by the Alzheimer's Association. As they are the main parties affected by the proposed decision, the perspectives of the patient and their families will be put into conversation.

Dementia is defined by the Alzheimer's Association, a leader in dementia advocacy, as "a general term for loss of memory, language, problem-solving and other thinking abilities that are severe enough to interfere with daily life" ("What is Dementia" *alz.org*). In the U.S. alone, 30% of people over 65 are dying from or with dementia. Although dementia is a terminal condition, it is very difficult to secure hospice care (life expectancy of 6 months or less) since the decline is "slow and unpredictable". According to Krista Harrison, a doctor who specializes in caring for terminally ill patients, the unpredictable nature of dementia's progression and prevention from access to proper care only adds to the pain and discomfort of "frequent burdensome hospitalizations, and invasive procedures that neither improve length nor quality of life". In addition, families suffer from the diagnosis, spending an average "\$61,000 out-of-pocket in the last 5 years of life – twice as much as other terminal conditions – in addition to thousands of hours caregiving that may negatively impact financial, physical, and mental wellbeing" (Harrison et al. 1548-1549).

MAiD has already been legalized in several countries across the world, including the Netherlands and Canada. In Canada, within the first 6 months after MAiD was legalized, 72 medically assisted deaths took place in Vancouver Island Health Authority (Robertson et al. 305). Though several countries have a terminal illness requirement, many do not (see table 1). This is because the structure of the law in those countries is based on moral principle. In the Netherlands, the requirement for MAiD is that “if a patient is suffering greatly from serious and irreversible illness, the patient would be allowed to die” (Lynn 17). Dementia surely qualifies as this, which is why it is legal in the Netherlands. One example of this was a case where a doctor was acquitted for giving MAiD to a woman with Alzheimer’s after her previous request. This woman had already met the requirements of ‘unbearable and endless suffering’ and approval of two doctors. After the case, a court ruled that a doctor “may carry out a written request

Table 1
Comparison of some international jurisdictions that allow MAiD (July 2022)^a

Jurisdiction	Terminal illness Requirement	Minors	Mental Disorder	Advance Requests	Methods of Provision
Netherlands	No	Yes*	Yes	Yes, 1,2*	Both
Belgium	No (yes for minors)	Yes*	Yes	Yes*	Euthanasia*
Luxembourg	No	No	Yes	Yes*	Both
Switzerland	No	No*	Yes	No	Assisted suicide
Spain	No	No	Yes*	Yes	Both
Canada	No	No	Yes, in March 2027	No	Both
Colombia	No*	Yes*	Yes*	Yes	Both
United States ²	Yes*	No	No	No	Assisted suicide
New Zealand	Yes	No	No	No	Both
Australia ³	Yes	No	No	No	Varies across states

Source: Doelle, Alida. "Medical Assistance in Dying around the World." *Dying with Dignity* Canada, 28 Oct. 2022, www.dyingwithdignity.ca/blog/medical-assistance-in-dying-around-the-world/. Accessed 24 Mar. 2025.

a. Note: This chart summarizes access to MAiD in each country under certain categories; content with an asterisk is more complex than the simplified chart can convey.

beforehand for euthanasia in people with advanced dementia” (qtd. in BBC *bbc.com*). This preserves a patient’s autonomy and also allows them to live longer than they might in other circumstances, but is illegal in the U.S..

This leaves patients that are truly suffering with no other option than to take matters into their own hands. Although suicide is legal, this is not a viable option for most terminally ill patients; some dementia patients are so weak that they can no longer commit suicide. This forces them to keep living, since there is no legal way for anyone to help them (Benatar 2). This could be a driving factor in the increased suicide rate of recently diagnosed dementia patients. A study

by the Alzheimer's Association found that almost half of dementia patient suicides occurred "within the first 90 days" after a dementia diagnosis. The most common method is by firearm(63.3%), which is messy and can involve law enforcement (Schmutte et al. 5-6).

This has a great impact on the family of the patient, as they will most likely miss their chance to say goodbye. A study done by the Journal of Pain and Symptom Management found that while 90% of family members wanted to be at their loved ones' side as they died, only 79% were able to. This deeply affects the grief of remaining family members and can lead to depression and more severe grief (Otani et al. 273). However, death by suicide is even more devastating for a partner. In a study by the BMC Psychiatry journal, MAiD "had a protective effect on the severity of the grief experiences of bereaved partners compared to suicide" and also protected them from society's stigma surrounding suicide. The study showed that a violent death is "strongly associated with difficulty in accepting the loss", PTSD, and depression (Snijdwind et al. 10). As the leading cause of dementia patient deaths is firearms, which is a violent death, and there is typically less than a 3-month period for the partner to come to terms with the diagnosis, let alone a suicide, this places a great strain on a partner.

On the other hand, MAiD allows a patient to end their life in a dignified way, surrounded by their family, and gives their loved ones time to say goodbye. This is illustrated in a CBS news report where Barabara Goodfriend went through her last day with her family and was sure of her choice to use MAiD after being diagnosed with ALS. When asked if she was ready, she replied "I'm so ready" (CBS 0:2:32-0:2:33). Goodfriend is not alone in her opinion, with a vote of 46 to 28 in the Dutch senate for the legalization of euthanasia and over 90% of the Netherlands supporting the bill (The Guardian *theguardian.com*). Even in the U.S., the Colorado MAiD law passed by a "30-point margin", which was "larger than any ballot initiative in Colorado history" and that received support from every area and demographic in Colorado (Lynn 33).

Although MAiD for dementia patients should be legal, it should not be required, or forced on anyone. Doctor Johannes Mulder PhD reminds that the patient can choose to call off the procedure at any time, same as with after-death organ donation ("Facilitating" 596). However, if the patient is committed to their choice of MAiD, they are a prime candidate for organ donation since the doctors can be ready to preserve their organs. In the U.S., over 100,000 people are on the national transplant waiting list and 17 people die each day waiting for an organ transplant (HRSA *organdonor.gov*). Although some might argue that MAiD could be used as an excuse to get the patient's organs, in other countries where simultaneous MAiD and organ donation are legal, this is not the case. In the Netherlands, it was shown that "the wish for organ donation after MAiD originates from patients themselves" and not the doctor or family. Organ donation can give the patients "a sense of purpose and fits with the autonomous determination of the final phase of their life" ("Facilitating" 595).

Procedures dictate that the doctor cannot mention organ donation at all, as it is completely separate from the MAiD proceedings. Similarly as to how the patient must request MAiD, the patient must express their desire to be an organ donor after death. However, the organ donation process typically takes place in a hospital, while the reason some patients request

MAiD is so that they do not have to enter the hospital. A new approach to MAiD and organ donation is necessary to solve this problem. Doctor Mulder reports that after using an anesthesia bridge from the home where the patient ‘died’, the doctors take the patient to the hospital and prepare their body for organ donation. This allows the organs to be healthy and not have gone a long time without blood. In fact, the first time this was done in the Netherlands, there was a national broadcast to accompany the event (“Organ Donation” 1305-1306). This has proven largely successful, as in 50 cases of MAiD and organ donation, “no issue has arisen regarding damage to organs as a consequence of the process of MAiD”. Though this does benefit other people, the patient undergoing MAiD and organ donation can have every confidence that “the interests of the living patient who has chosen MAiD always take priority over whatever organs they might pass on after death” (“Facilitating” 596).

In addition to preserving the autonomy of patients, MAiD can ease the burden that caring for patients places on their families. Alzheimer’s, the most common form of dementia (“What is Dementia” *alz.org*), affects an estimated 6.9 million patients. In addition, “More than 11 million Americans provide unpaid care for a family member or friend with dementia” which is valued at \$350 billion. It is predicted that this number will drastically increase to 13.8 million patients over 65 by 2060 (“Facts and Figures” *alz.org*). The burden that unpaid care places on the family can strain the relationships between the family and patient.

In some instances, regardless of all of the signs that the patient wishes to and is fully aware of dying, some family members cannot let them. An online forum for aging care discusses this, with one user posting, “Physical pain, a loved one not letting go and the dying person feels they have to stay alive for that person. It's selfish to want someone to be stuck suffering in a body that's torturing them just so they won't have to face the loss” (Nakeeta 2021). MAiD gives the family a chance to prepare themselves and recognize that their loved one wants to move on. However, without their support, the patient could force themselves to hold on. The family could also be holding on to an idea of peace from earlier times, often forgetting the drawbacks that were present. As a European historian explains, nostalgia can idealize history until it is distorted, as memory is “notoriously unreliable” (Norberg 13). Families can idolize their memories of a ‘simpler’ time and childhood and relate the memories to the patient, especially if they are the child of the patient.

There are wide concerns about those using MAiD being uneducated or desperate; however, a study by a PhD professor has shown that this is not true. Most patients that died by MAiD between 1998-2020 were elderly with a degree of college education. Almost half of these patients were married and died at home (Kozlov et al. 3041). These statistics show that MAiD is not being abused by under-educated, poor, or suicidal people but is an autonomous choice by an individual who is ready to die. In the end, MAiD is an option, not a choice, and there is legislation around the globe to prevent a patient from being forced into it. Having MAiD as a ‘safety net’ could greatly increase the quality of life as patients approach the end. In fact, according to Death with Dignity, 30% percent of patients who received a prescription never took it (Haring *deathwithdignity.org*).

Another prevailing concern about MAiD is the effect on doctors, who have dedicated their lives to saving lives and are now being asked to end them. All licensed doctors can prescribe lethal medications; however, they all have the right to refuse to participate (Haring *deathwithdignity.org*). This ensures that both parties, the patient and doctor, are willing to perform the procedure. Worldwide, there are many organizations that advocate for the autonomous option of MAiD, including Death With Dignity and World Federation Right to Die Societies. Petitioning for MAiD is an exercise of human rights on behalf of those who cannot help themselves. Let's fight to restore autonomy to those who have lost it.

Works Cited

- Aging Care. "What Are the Reason Why the Dying Can't Let Go?" Agingcare.com, 9 May 2020, <https://www.agingcare.com/questions/what-are-the-reason-why-the-dying-cant-let-go-470453.htm>. Accessed 20 Mar. 2025.
- Alzheimer's Association. "2024 Alzheimer's Disease Facts and Figures. Special Report: Mapping a Better Future for Dementia Care Navigation." Alzheimer's Association, 2024.
- Alzheimer's Association. "What Is Dementia?" *Alzheimer's Disease and Dementia*, Alzheimer's Association, 2025, www.alz.org/alzheimers-dementia/what-is-dementia. Accessed 11 Mar. 2025.
- BBC. "Dutch Court Expands Euthanasia Law on Dementia." BBC News, 21 Apr. 2020, www.bbc.com/news/world-europe-52367644. Accessed 11 Mar. 2025.
- Benatar, David. "Should There Be a Legal Right to Die?" *Current Oncology* (Toronto, Ont.), vol. 17, no. 5, 2010, pp. 2–3, www.ncbi.nlm.nih.gov/pmc/articles/PMC2949364/, <https://doi.org/10.3747/co.v17i5.671>. Accessed 5 Mar. 2025.
- CBS Evening News. "A Woman with ALS Shares Her Last Day after Choosing "Medical Aid in Dying."" YouTube, 31 Jan. 2025, www.youtube.com/watch?v=jeV3KUoW31U. Accessed 8 Mar. 2025.
- Compassion & Choices. "States Where Medical Aid in Dying Is Authorized - Compassion & Choices." Compassion & Choices, 2015, compassionandchoices.org/states-where-medical-aid-in-dying-is-authorized/. Accessed 11 Mar. 2025.
- Cornell Law School. "Suicide." LII / Legal Information Institute, Aug. 2021, www.law.cornell.edu/wex/suicide. Accessed 11 Mar. 2025.
- Haring, Chris. "Medical Aid in Dying End-of-Life Option Offers Dignity." *Death with Dignity*, Death With Dignity, 29 Mar. 2023, deathwithdignity.org/news/2023/03/3-29-23-senior-guide/. Accessed 23 Mar. 2025.
- Harrison, Krista L., et al. "Dying with Dementia: Under-Recognized and Stigmatized." *Journal of the American Geriatrics Society*, vol. 67, no. 8, 1 Aug. 2019, pp. 1548–1551, www.ncbi.nlm.nih.gov/pmc/articles/PMC6684346/, <https://doi.org/10.1111/jgs.15895>. Accessed 13 Mar. 2025.
- HRSA. "Organ Donation Statistics." *Www.organdonor.gov*, Health Resources & Services Administration, Oct. 2024, www.organdonor.gov/learn/organ-donation-statistics. Accessed 13 Mar. 2025.
- Kozlov, Elissa, et al. "Aggregating 23 Years of Data on Medical Aid in Dying in the United States." *Journal of the American Geriatrics Society*, vol. 70, no. 10, 16 June 2022, pp. 3040–3044, agsjournals.onlinelibrary.wiley.com/doi/full/10.1111/jgs.17925, <https://doi.org/10.1111/jgs.17925>. Accessed 20 Mar. 2025.
- Lynn, Joanne. *Physician-Assisted Death: Scanning the Landscape: Proceedings of a Workshop*. 2018. Nih.gov, National Academies Press (US), 27 June 2019, pp. 9–33, www.ncbi.nlm.nih.gov/books/NBK525943/. Accessed 7 Mar. 2025.

- “Medical Aid in Dying Is Not Assisted Suicide, Suicide or Euthanasia - Compassion and Choices.” <https://Compassionandchoices.org/compassionandchoices.org/resource/not-assisted-suicide/>. Accessed 7 Mar. 2025.
- Mishara, Brian L., and David N. Weisstub. “The Legal Status of Suicide: A Global Review.” *International Journal of Law and Psychiatry*, vol. 44, Jan. 2016, pp. 54–74, www.sciencedirect.com/science/article/abs/pii/S0160252715001429, <https://doi.org/10.1016/j.ijlp.2015.08.032>. Accessed 7 Mar. 2025.
- Mulder, Johannes. “Facilitating the Wishes of Patients Who Choose Both MAiD and Organ Donation.” *Canadian Medical Association Journal*, vol. 191, no. 22, 2 June 2019, pp. E595–E596, pmc.ncbi.nlm.nih.gov/articles/PMC6546575/#b10-191e595, <https://doi.org/10.1503/cmaj.190352>. Accessed 12 July 2020.
- Mulder, Johannes, and Johan P.C. Sonneveld. “Organ Donation after Medical Assistance in Dying at Home.” *Canadian Medical Association Journal*, vol. 190, no. 44, 4 Nov. 2018, pp. E1305–E1306, www.cmaj.ca/content/190/44/E1305, <https://doi.org/10.1503/cmaj.170517>. Accessed 6 Mar. 2025.
- Norberg, Johan. “False Nostalgia.” *Reason Magazine*, Jan. 2022, pp. 8–13.
- Otani, Hiroyuki, et al. “Meaningful Communication before Death, but Not Present at the Time of Death Itself, Is Associated with Better Outcomes on Measures of Depression and Complicated Grief among Bereaved Family Members of Cancer Patients.” *Journal of Pain and Symptom Management*, vol. 54, no. 3, Sept. 2017, pp. 273–279, <https://doi.org/10.1016/j.jpainsymman.2017.07.010>.
- Robertson, W. D., et al. “Case Review of Medically Assisted Deaths on Vancouver Island: A Total of 72 Assisted Deaths Took Place in the Vancouver Island Health Authority in the First 6 Months after Medical Assistance in Dying Ceased to Be Illegal in Canada.” *British Columbia Medical Journal*, vol. 59, no. 6, July 2017, pp. 305–09. EBSCOhost, research.ebsco.com/linkprocessor/plink?id=f40abdb8-7de9-3149-b1e8-f8ce06c42a62.
- Schmutte, Timothy, et al. “Suicide Risk in First Year after Dementia Diagnosis in Older Adults.” *Alzheimer’s & Dementia*, 25 May 2021, <https://pmc.ncbi.nlm.nih.gov/articles/PMC8613307/>. Accessed 20 Mar. 2025.
- Snijdewind, M. C., et al. “‘I Lost so Much More than My Partner’ – Bereaved Partners’ Grief Experiences Following Suicide or Physician-Assisted Dying in Case of a Mental Disorder.” *BMC Psychiatry*, vol. 22, no. 1, July 2022, pp. 1–13. EBSCOhost, <https://doi.org/10.1186/s12888-022-04098-5>.
- The Guardian. “Dutch Pass Euthanasia Bill.” *The Guardian*, 11 Apr. 2001, www.theguardian.com/world/2001/apr/11/andrewosborn.sarahboseley. Accessed 13 Mar. 2025.
- United Nations. “Universal Declaration of Human Rights.” United Nations, United Nations, 10 Dec. 1948, www.un.org/en/about-us/universal-declaration-of-human-rights. Accessed 7 Mar. 2025.

Rescuing Wall Street While Overlooking Main Street?

A Comparative Analysis of the 2008 Financial Crisis Response By Krishna Rathuryan

Abstract

The 2008 Financial Crisis had severe economic and psychological impacts on low to middle income households, including rising unemployment, wealth loss, and increased financial stress. This paper explores how government responses to the 2008 Crisis, such as the Economic Stimulus Act of 2008, compared to those of another major American financial crisis: the Great Depression. By analyzing the historical policies and the broader economic environment during these periods, this paper identifies the common challenges faced by low and middle income households and evaluates whether the policies introduced in 2008 were, compared to the Great Depression, as effective and sufficient for mitigating the negative impacts on these households. The findings suggest that while government initiatives were taken during both crises, the government policy for mitigating the negative impacts on low to middle income households during the 2008 Crisis was not as effective compared to that of during the Great Depression.

Introduction

A financial crisis is a severe disruption in the economy that affects the stability of financial institutions, markets, and systems, often resulting in widespread bankruptcies, asset devaluation, and a loss of public confidence (Kenton). These crises are typically marked by a sudden and sharp decline in economic activity, rising unemployment, and an erosion of household wealth and income. The 2008 financial crisis, which stemmed from things like subprime mortgages, lack of regulation, and the subsequent collapse of the U.S. housing market, is one of the most significant economic events in recent history (Weinberg). While the financial sector faced a devastating collapse, the impacts on low to middle income households were equally severe, as rising unemployment, foreclosures, and financial stress took a significant toll on their wellbeing.

Government policy and responses to such crises are crucial in mitigating the negative social and economic consequences, including on low to middle households. This paper compares the U.S. response to the 2008 crisis, also known as the Great Recession, with past interventions during the Great Depression—another large American financial crisis—such as the New Deal when it comes to mitigating the negative impacts on low to middle income households. By doing so, it specifically evaluates how well the Great Recession policies addressed the needs and problems of low and middle income families and whether they were as successful and sufficient as those during the Great Depression. By examining both crises, it was found that the Great Depression policies were better and more effective than those introduced during the Great Recession.

The Two Crises

The Great Depression (1929-1939), had multiple causes, both immediate and systemic. One of the most significant triggers was the stock market crash of October 1929, when speculative investments led to an unsustainable bubble that burst, wiping out millions of dollars in wealth (“Causes”). However, the crash alone did not cause the Depression; overproduction in agriculture and manufacturing led to surpluses, which drove prices down and resulted in many businesses and individuals going bankrupt (“Economic Hardship”). Legislation like the Smoot-Hawley Tariff Act of 1930 added salt to the wound by drastically reducing international trade (“Causes”). Adding on, banks had been poorly regulated, and many failed when depositors rushed to withdraw their funds in panic, further destabilizing the financial system (“Causes”). The Depression led to widespread unemployment, poverty, and a collapse of the financial system, as businesses closed, and banks failed.

The Great Recession (2007-2009), however, was primarily caused by a combination of risky financial practices, lax regulation, and a housing bubble. At the center of the crisis were subprime mortgages—high-risk loans given to borrowers with poor credit histories (Coghlan). They were bundled into complex financial products called MBSs—mortgage-backed securities—and CDOs—collateralized debt obligation—and they were sold to banks and investors (Coghlan). Simply put, these products pooled together various debts, including these risky house mortgages. Over time, these investors could expect to get a share of the interest payments on these debts and for the value of their CDOs and MBSs to go up. Prior to the crisis, home prices, fueled by giving big loans to “unworthy” borrowers, also known as subprime mortgages, and low interest rates, had been booming (“What”). So, when home prices started to fall in 2007, these “unworthy” borrowers began to owe more than what their home itself was worth, leading to many foreclosures and loan defaults (Gratton). This led to the value of these CDOs and MBSs to plummet, leading to widespread financial instability, the collapse of major financial institutions, and high unemployment, among others.

Impacts on Low to Middle Income Households

The Great Depression devastated low and middle-income families in the United States, causing widespread economic hardship and psychological distress. Unemployment had skyrocketed to 25% by 1933, with some urban areas experiencing even higher rates (“Bureau”). Average family income plummeted by 40% between 1929 and 1933, from \$2,300 to \$1,500 (“Children”). Rural communities were particularly hard-hit, with 60% of rural households and 82% of farm families classified as “impoverished” (“Everyday Life”). Food insecurity also increased, and psychological impacts were severe, as many felt worthless, anxious, and stressed. Suicide rates rose from 18 per 100,000 in 1928 to 22 per 100,000 in 1932 (“Suicides”). The Great Recession significantly impacted low and middle-income households as well. The average duration of unemployment increased a lot, and the unemployment rate peaked at 10% in 2009, leaving 15 million Americans out of work (“Great Recession”).

In addition, median household income fell by 7.7% between 2007 and 2010, from \$49,600 to \$45,800 (Peterson). About 3.8 million foreclosures were completed between 2007

and 2010, and by the end of 2011, 31.4% of homeowners were “underwater” on their mortgages (“Zillow”). The net worth of American households fell from \$69 trillion in 2007 to \$55 trillion in 2009 (Rich). Poverty rates increased from 12.5% in 2007 to 15.1% in 2010 as well, adding 6.3 million people to the ranks of the poor (McCorkell). Food insecurity increased, with the number of Americans receiving food stamps and benefits rising from 26.3 million in 2007 to 40.3 million by 2010 (“SNAP”). Furthermore, the psychological impacts were significant, with a 6.5% increase in men’s suicide rates in 2009, attributable to the recession (Singh).

Government Policies

In order to combat the impacts on low to middle income households during the Great Depression, the government, especially under Franklin D. Roosevelt’s presidency, brought about many changes. First off, as part of the New Deal, the government introduced the National Youth Administration, Works Progress Administration, and Civilian Conservation Corps, which are generally shortened to NYA, WPA, and CCC, respectively (“New Deal”). The CCC, WPA, and NYA not only helped give jobs and a livable wage to millions of unemployed Americans, but it also provided them with other benefits such as healthcare, housing, food, and additional financial assistance (“New Deal”). In addition, these initiatives, with the addition of millions of new jobs, helped stimulate economic growth during the Great Depression. The Social Security Act of 1935 provided pensions for the elderly and gave unemployment insurance, while the National Industrial Recovery Act (NIRA) established fair wages and prices in different industries (“New Deal Programs”). The Federal Emergency Relief Administration (FERA) was created as well, which added tens of millions of unskilled jobs that anyone could do (“New Deal Programs”). While adding jobs to the market is more expensive than direct cash support, it is psychologically more beneficial for the unemployed, since having a job can boost self-esteem a lot (Botes). As part of the FERA, the government, which had set aside \$3.1 billion (~\$70.8 billion in 2024), had actually also given direct cash support to low and middle income families for basic necessities like food and shelter (“Family Experiences”). Numerous food stamps were given to needy families as well.

During the Great Recession, on the other hand, the focus for mitigating the impacts on low to middle income households was more on direct cash support rather than new employment opportunities. The Economic Stimulus Act of 2008 gave most of these households up to a \$1200 one-time payment, with an additional \$300 per child, and it helped approximately 130 million households fund their basic necessities (“President Bush”; Zerwitz). The Supplemental Nutrition Assistance Program (SNAP) reduced food insecurity and eased the problem of hunger for millions of Americans (Nord and Prell). In 2009, the American Recovery and Reinvestment Act (ARRA) set aside \$831 billion to invest and give tax cuts to businesses and households, with each worker getting \$400 (Hayes). It also expanded unemployment benefits and increased the amount of weeks one could be unemployed and still receive benefits (Hayes). The Emergency Unemployment Compensation Act of 2008 (EUC08) provided more temporary unemployment benefits to those who had exhausted their regular unemployment benefits (“Emergency”). The

Home Affordable Modification Program (HAMP) helped low and middle income families homeowners at risk of foreclosure by making mortgage payments more affordable (“Home”). To give credit where it’s due, however, programs and policies like the ARRA and SNAP, did help create a lot of jobs by encouraging consumer spending and helping businesses through investment/tax cuts, but they were nowhere as significant as those during the Great Depression.

The Effective Approach

The Great Depression lasted for a lot longer and had much worse impacts on low to middle income households than the Great Recession. For example, the unemployment rate during the Great Depression was 2.5 times greater than that of during the Great Recession (“Bureau”; “Great Recession”). However, despite that, it seems to be that the government response during the Great Depression, when it comes to mitigating the negative impacts on low to middle income households, was more effective and successful than that of during the Great Recession. One of the most pressing issues during both crises was unemployment. The peak unemployment rate during the Great Recession was 10% in 2009, almost twice the pre-crisis unemployment rate, while the peak unemployment rate during the Great Depression was 24.9% in 1933, almost eight times the pre-crisis rate (“Bureau”; “Great Recession”). When looking at the changes in unemployment rates during and after the crises, it can be seen that the policies and programs of the Great Depression did a better job of handling unemployment than those during the Great Recession. By 1937, just four years after the peak unemployment rate in 1933, the unemployment rate had dropped by a staggering ~12% (“Bureau”). This period of time (1933-1937) also coincides exactly with the introduction of the New Deal and its various programs, so the unemployment drop is most likely attributable to that, considering that the New Deal added millions of new jobs. Unemployment did slowly creep up in the late 1930s, but it dropped back with the introduction of more policies and World War II (Waiwood). Meanwhile, unemployment rates during the Great Recession would continue to climb and wouldn’t start to drop significantly until the early 2010s (post-recession), showing that the policies and programs introduced did not effectively handle unemployment during the crisis (“Great Recession”).

The next major issue, at least concerning low and middle income households, was wealth loss (falling household income) and negative psychological effects. During the Great Depression, much more aggressive and dynamic actions were taken to combat these issues. As mentioned before, the various programs of the New Deal, such as the CCC and WPA, helped pump tens of millions of new jobs into the market, which not only gave stable income and wealth to low and middle income households but also psychologically uplifted them by giving them more self-esteem and confidence with their new jobs (Botes; “New Deal”). By 1935, the median household income would start recovering substantially, easing wealth loss (Wattenberg). Programs like the FERA also gave direct cash support to struggling families as a second layer of aid. During the Great Recession, on the other hand, the median household income would not start recovering substantially until 2014, almost five years after the crisis itself (Guzman). Policies like the ARRA didn’t create or save sufficient jobs, as seen by the fact that unemployment

continued to rise or plateau during and after the crisis (“Great Recession”); many low and middle income workers were still left unemployed, thus resulting in many negative psychological impacts such as low self-esteem. There was a lot of aid and unemployment assistance given through programs like the HAMP and EUC08, but it wasn’t even close for families to feel relief and recover from wealth loss, at least in the long-term. Most low and middle income families received a measly \$1000 or \$2000, which is barely enough to cover for a month’s expenses (Hayes; Zerwitz). Also, the top 1% of earners captured 95% of income gains in the first three years of recovery after the crisis, further showing the ineffectiveness of the Great Recession response in reducing wealth loss in low to middle income households (Weber).

In general, the policy actions taken during the Great Depression were more effective and sufficient than those taken during the Great Recession when it comes to mitigating the negative impacts on low to middle income households, including unemployment, wealth loss, and psychological distress. As seen clearly, the programs and policies introduced during the Great Depression were much more aggressive and long-term oriented (creating jobs) compared to those of the Great Recession, which were more “short-term oriented (mainly direct cash support). This is mainly the reason why the Great Depression response was more successful in reducing the negative effects on low to middle income households. Short-term solutions like cash payments only give temporary relief and don’t actually tackle the issues in the long run. According to Matthew Wilson, who is from a middle income family that went through the Great Recession, “what unemployment paid was much less than what he [his dad] was normally making weekly. My family had bills that needed to be paid...unemployment helped, but it was a temporary fix” (Wilson). On the other hand, long-term solutions like creating more jobs in the market are more sustainable, and they not only help the economy but also give true relief for struggling households through the form of stable income and benefits. Thus, the impacts on low to middle income households during the Great Recession could have been handled better if the government and policymakers had focused a lot more on long-term solutions like job creation rather than mostly on short-term solutions like direct cash support. Looking ahead, this same approach would likely be the most effective for mitigating the negative impacts on low to middle income households during future crises as well.

Conclusion

The policies and programs introduced during the Great Depression seemed to be much better and more effective than those introduced during the Great Recession. Despite the Great Depression overall being a larger crisis than the Great Recession, with more unemployment, wealth loss, etc., and thus likely harder to handle, the Great Depression saw much better and drastic improvements when it came to these issues. New Deal programs like the CCC and FERA created long-lasting safety nets and really helped fix the root causes of the issues faced by low to middle income households. Meanwhile, the Great Recession response was more short-term, and any policies introduced for mitigating the negative impacts on low to middle income households, like the Economic Stimulus Act of 2008 and EUC08, were not as effective and successful.

Works Cited

- Botes, Alice. "What Psychological Benefits Do You Get from Work?" *Psychology Today*, Sussex Publishers, 22 Aug. 2019, www.psychologytoday.com/blog/in-practice/201908.
- "Bureau of Labor Statistics Data." U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, data.bls.gov/timeseries/LFU21000100&series_id=LFU2200
- "Causes of the Great Depression." *Encyclopædia Britannica*, Encyclopædia Britannica, www.britannica.com/story/causes-of-the-great-depression.
- "Children & The Great Depression." *Digital History*, www.digitalhistory.uh.edu/active_learning/explorations/children_depression.
- Coghlan, Erin, et al. "What Really Caused the Great Recession?" *Institute for Research on Labor and Employment*, 18 July 2022, irle.berkeley.edu/publications/irle-policy-brief/the-great-recession.
- "Economic Hardship and the Great Depression: Oklahoma Historical Society." *Oklahoma Historical Society*, www.okhistory.org/learn/tenant-farming8.
- "Emergency Unemployment Compensation 2008 (EUC) Program." U.S. Department of Labor, oui.doleta.gov/unemploy/pdf/euc08.pdf.
- "Everyday Life During The Great Depression." *Lumen*, courses.lumenlearning.com/wm-ushistory2/chapter/everyday-life.
- "Family Experiences and New Deal Relief." *National Archives and Records Administration*, <https://www.archives.gov/publications/prologue/2012/fall/fera.html#>.
- Gratton, Peter. "The Stock Market Crash of 2008." *Investopedia*, www.investopedia.com/articles/economics/09/subprime-market-2008.
- "Great Recession, Great Recovery? Trends from the Current Population Survey." *Monthly Labor Review*, U.S. Bureau of Labor Statistics, www.bls.gov/opub/mlr/2018/article/great-recession.
- Guzman, Gloria. "New Data Show Income Increased in 14 States and 10 of the Largest Metros." *Census.Gov*, 3 Oct. 2024, www.census.gov/library/stories/2019/09/us-median-household-income.
- Hayes, Adam. "American Recovery and Reinvestment Act (ARRA): Objectives and Faqs." *Investopedia*, www.investopedia.com/terms/a/american-recovery-and-reinvestment-act.asp, 25 June 2023.
- "Home Affordable Modification Program (HAMP)." U.S. Department of the Treasury, home.treasury.gov/data/troubled-assets-relief-program/housing/mha/hamp.
- Kenton, Will. "Financial Crisis: Definition, Causes, and Examples." *Investopedia*, www.investopedia.com/terms/f/financial-crisis.asp.
- McCorkell, Lisa, and Sara Hinkley. "The Great Recession, Families, and the Safety Net." *Institute for Research on Labor and Employment*, 18 July 2022, irle.berkeley.edu/publications/irle-policy-brief/the-great-recession-families.

“New Deal.” Encyclopædia Britannica, Encyclopædia Britannica, 28 Oct. 2024,
www.britannica.com/event/New-Deal.

“New Deal Programs.” Living New Deal, 16 Jan. 2024,
<https://livingnewdeal.org/history-of-the-new-deal/programs/>.

Nord, Mark, and Mark Prell. “Food Security of SNAP Recipients Improved Following the 2009 Stimulus Package.” USDA ERS, USDA, 13 June 2011,
[www.ers.usda.gov/amber-waves/2011/june/food-security-of-snap/#:~:text=SNAP%](http://www.ers.usda.gov/amber-waves/2011/june/food-security-of-snap/#:~:text=SNAP%20).

Peterson, Kristina. “Family Net Worth Fell Almost 40% between 2007-2010.” Wall Street Journal, 11 June 2012, www.wsj.com/articles/BL-REB-16496.

Rich, Robert. “The Great Recession.” Federal Reserve History,
www.federalreservehistory.org/essays/great-recession-of-200709.

Singh, Maanvi. “Suicide Rate in U.S. and Europe Climbed during Great Recession.” NPR, NPR, 12 June 2014, www.npr.org/sections/krulwich/2014/06/11/318885533/suicide-rate-in-u-s.

“SNAP Participation Response to 2008-10 Economic Decline Was Similar to Previous Downturns.” USDA ERS - Chart Detail,
www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?817.

“Suicides Go Up When Economy Goes Down.” Tristate Homepage, 4 Nov. 2011,
www.tristatehomepage.com/news/suicides-go-up-when-economy-goes-down/.

Waiwood, Patricia. “Recession of 1937–38.” Federal Reserve History,
www.federalreservehistory.org/essays/recession-of-1937-38.

Wattenberg, Ben. “Money.” The First Measured Century, PBS, <https://www.pbs.org/fmc/book/9money3.htm>.

Weber, Peter. “Charts: How the Rich Won the Great Recession”, The Week, 11 Jan. 2015,
theweek.com/articles/460179/charts-how-rich-won-great-recession#:~:text=And%20beca use%the%20income%20of,wealth%20during%20the%20economic%20recovery.

Weinberg, John. “The Great Recession and Its Aftermath.” Federal Reserve History,
www.federalreservehistory.org/essays/great-recession.

“What Caused the Last Housing Boom?” World Economic Forum,
www.weforum.org/stories/2015/02/what-caused-the-last-housing-boom/.

Wilson, Matthew. “My Dad Lost His Job during the Great Recession in 2008. Here’s What Life Was like and How My Family Made It Through.” Business Insider, 28 Apr. 2020,
www.businessinsider.com/i-was-13-2008-recession-dad-lost-job-tips-2020.

Zerwitz, Donna. “Did the 2008 Tax Rebates Stimulate Spending?” NBER,
www.nber.org/digest/mar09/did-2008-tax-rebates-stimulate-spending.

“Zillow Negative Equity Report.” Mediaroom,
filecache.mediaroom.com/mr5mr_zillow/203016/download.

Global Trade and Domestic Disparities: How U.S. Policy Shapes South Korean Inequality

By Jason Chae

Abstract

This study examines the relationship between U.S. tariff increases and income inequality in South Korea. Using data from 1980 to 2025 and an Ordinary Least Squares (OLS) regression GDP per capita, foreign direct investment (FDI), trade openness, a dummy variable for Donald Trump's presidency, and U.S. average tariff rates as independent variables, with two separate measures of inequality as dependent variables, this paper finds that higher U.S. tariffs are associated with a reduction in income inequality. Through the regression, the effects that each of the independent variables has on openness is also explored.

Introduction

In recent years, the world has witnessed the re-emergence of old protectionist trade policies that were once popular in the early 1900s. Specifically, Donald Trump's former and current administration have both threatened and implemented significant tariffs, including proposals as high as 200% on alcohol from the EU (Wile et al.). South Korea is one of the countries that have been threatened with such tariffs, causing concern within the country and around the world about the economic fallout that may result from them. However, despite the significant focus placed on the tariffs themselves, there is currently a lack of research on how they might impact other societal or economic indicators, specifically on how these tariffs could affect income inequality. In such a heavily export-driven nation as South Korea, where tariffs in the triple digits are being threatened, what is the impact on income inequality that these tariffs may have?

By looking at previous research on the interaction between trade liberalization, openness, and income inequality, while running an Ordinary Least Squares (OLS) regression with measurements of income inequality as dependent variables, we can determine that the U.S. tariff increases threatened by the Trump administration can cause decreasing income inequality in South Korea. However, this reduction in income inequality, accompanied by economic decline or stagnation, threatens to be a bane rather than a boon for the low-income workers in South Korea.

Literature Review

Overview:

This segment of the paper will examine prior economic research concerning the effect that varying trade policies has on income inequality, the current causes of income inequality in South Korea, and the effect that Trump's protectionist policies have had on this country.

Global Insights on Inequality

Trade Liberalization:

Prior studies that utilize global data sources confirm a positive correlation between trade liberalization and increased income inequality when looking through the perspective of import

competition. The underlying reasoning behind this correlation is that trade liberalization, through the form of lower tariffs or the removal of quotas for imports, can increase the competitiveness of industries due to the arrival of a larger quantity of cheap imported goods. Through an increase in domestic competition, several middle-class jobs are erased, especially with the depletion in middle-tier firms which aren't able to keep up with the increased competition. Meanwhile, high-class jobs benefit due to top-tier firms being better able to penetrate foreign markets with trade liberalization while low-class jobs benefit due to the decrease in price of goods through import competition. Thus, as these middle-class jobs are erased, the newly educated yet unemployed workers move either to low-class or high-class jobs, increasing overall job polarization and in turn, income inequality (Furusawa et al. 21-23).

An example of this correlation between trade liberalization and income inequality is Keller and Utar's observations when tracking the effects of Denmark removing its trade restrictions targeting Chinese textile exports. With the implementation of trade liberalization within the textile industry in Denmark, a significantly higher rate of middle-wage job losses or job transfers could be observed in the industry exposed to import competition compared to other unexposed industries. Additionally, a much larger rate of middle-wage personnel from the textile industry was seen transferring to either lower-wage jobs or higher-wage jobs when compared to personnel from industries unexposed to import competition (Keller and Utar 41-42). Hence, these findings suggest that from a global perspective, trade liberalization may have a serious effect on increasing job polarization and income inequality because of import competition.

Openness

When looking through a global perspective, researchers are currently unsure of the effect that openness has on income inequality. While some have come to the conclusion that an increase in a country's openness leads to an increase in income inequality, others have discovered that protectionist trade policies that decrease openness, such as tariffs or non-tariff restrictions, increase income inequality.

Through observing the economic indicators of 60 different countries, Aradhyula, Rahman, and Seenivasan were able to come to the conclusion that trade openness, when considering country geology, trade policies, corruption levels, and economic conditions, led to an overall increase in per-capita income as well as income inequality. In addition, a higher level of income inequality caused by trade openness was observed in developing countries when compared to developed countries (Aradhyula et al. 36).

Although similar, Savvides arrives at a slightly different conclusion regarding the effect of trade policies on income inequality. When using a recently compiled index of trade protection along with new data on income distribution for both developed and developing countries, a correlation between openness and increased income inequality was found for less developed countries. However, for developed countries, there was no effect on income inequality found due to trade policies (Savvides 371-372).

On the other hand, when analyzing pro-trade policies throughout Latin American countries in the late 1990s as well as declines in economic inequality during the early 2000s, Rosenfield, Mota, and Pereira postulated that trade openness had an effect of decreasing income inequality. Specifically, when observing the results gathered from his model, a significant decrease of 0.03 to 0.05 points in the Gini Index was found for every 1% increase in trade openness (as a percentage of GDP). Additionally, the effect of economic simplification, where low knowledge intensity products are both mainly produced and exported internationally, was found to decrease income inequality. However, as economic simplification is very dependent on commodity prices and domestic conditions, it carries risks when considering long-term economic growth and income distribution (Rosenfield et al. 486-487).

Korean Insights on Inequality

Trade Liberalization

Approaching the issue of tariff's effects on income inequality through a South Korean lens reveals further vigorous debate. While some researchers have found that trade liberalization increases income inequality in Korea, other researchers have found contradictory results suggesting that trade liberalization decreases income inequality in the country. However, there is no overall consensus on this topic and the effects of trade liberalization on income inequality remain cloudy.

Research from Lee attempted to find the relationship between international trade and wage inequality, specifically wage dispersion amongst workers in the same industry. Through data analysis of South Korean economic metrics and reasoning, it was discovered that increased openness corresponded to a rise in within-sector wage inequality. The reasoning behind this is that those engaged in export related industries face a higher skill premium due to increased competition in the quality of products abroad. A higher skill premium can cause the gap between the wages of skilled and unskilled workers to widen disproportionately, especially in the case of South Korea, where a large portion of economic activity is comprised of exports. Along with a heavy reliance on the technological sector in exports, where the skill of an employee is extremely important to the quality of exported goods, the reasoning behind Lee's research suggests that trade liberalization increases income inequality in South Korea (Lee 44-46).

Meanwhile, research from Cheong and Jung searches for the relationship between international trade and the difference in wages between unskilled and skilled workers. Through the use of the difference in differences method and the traditional wage equation—along with data examining the effects that free trade agreements between South Korea and the U.S. and EU had on the South Korean economy—researchers examined the consequences of trade liberalization on South Korean inequality, especially with regards to unskilled and skilled workers. The data revealed that after the free trade agreement was signed between the U.S., EU, and South Korea, wages of unskilled workers rose in an un-uniform manner. Thus, the research suggests that while trade liberalization may have had a negative effect on income inequality, the results are not fully

clear and the increases in wages may signal a shift in overall inequality rather than a decrease in inequality (Cheong and Jung 1, 12-13).

Global Connections To South Korea

Connecting South Korean specific data with global research on trade, economic growth, and income inequality can uncover new and interesting information regarding the causes of income inequality in South Korea.

By running multiple regression analyses on South Korean data with a dummy variable indicating Trump's presidency, Jung and Mun discovered that foreign ownership/foreign direct investment (FDI) increased in South Korea prior to their initial hypothesis (Jung and Mun 93-94). A correlation between Trump's presidency and an increase of FDI in South Korea is important because research from Wang, Park, and Choi suggests that countries with a higher GDP and FDI experience higher income inequality (Wang et al. 31). While this may create the perception that Trump's presidency will increase income inequality due to the increase of FDI in South Korea, Jung and Mun found that FDI was only able to increase due to South Korean companies increasing the amount of information released in order to account for lower profitability under Trump's protectionist policies. Since it is unknown whether South Korean companies will further increase the amount of information shared with foreign investors during Trump's second term, it is unclear from these sources whether Trump's second term will either increase or decrease income inequality.

Traditional economic concepts such as the Kuznets Curve also point to a decrease in income inequality as per capita GDP increases after reaching a certain turning point of economic development. This turning point can typically be classified as the transition from a developing to a developed economy due to the traits of developed economies such as deindustrialization, an increase in the service industry, and the increase of government assistance programs contributing to a more equitable income distribution. As Korea has seen a shift towards deindustrialization and an increase in the service sector in the past decades, along with massive increases in government welfare expenditure, South Korea is poised to see a decrease in income inequality should GDP per capita growth be observed (Hong 253)(Yoon).

Research from Shin and Kong also utilized South Korean-specific data, culture, and recent events to uncover the key factors that contributed to South Korean household inequality. While factors like aging population and the rise of single-person households have been linked to increased household inequality, it was discovered that one of the underlying causes of South Korean inequality was unemployment, especially during the 1997 and 2008 financial crises, which heavily impacted the country (Shin and Kong 34-37). However, Furusawa, Kellar, and Utar's research found that policies that promote trade liberalization increase the amount of unemployment. Therefore, it can be reasoned that if South Korea were to promote trade liberalization policies, unemployment could rise, thereby increasing inequality.

Recap

Overall, for global data sets, while literature confirms a positive correlation between trade liberalization and increased income inequality, the effect openness has on income inequality is still a heavily debated topic amongst researchers. When focusing on Korean-specific data, the debate only becomes more heated with contradicting effects of trade liberalization on income inequality populating South Korean literature. Thus, it seems that there is presently ongoing debate on the exact effects that trade liberalization, openness, and tariffs will have on income inequality amongst general research on the intersection between trade and inequality.

Methodology

To investigate what affects U.S. trade policy, specifically Donald Trump's pro-tariff policies, have on Korean income inequality, we utilized an OLS linear regression with data compiled from the Federal Reserve, World Bank, OECD, and other macroeconomic datasets.

Data

GINI Index (without Tariff Rate):

$$GiniIndex_{it} = \alpha + \beta_1 \log(GDPperCapita_{it}) + \beta_2 FDI_{it} + \beta_3 Openness_{it} + \beta_4 DummyTrump_{it} + \varepsilon_{it}$$

GINI Index (with Tariff Rate):

$$GiniIndex_{it} = \alpha + \beta_1 \cdot \log(GDPperCapita_{it}) + \beta_2 \cdot FDI_{it} + \beta_3 \cdot Openness_{it} + \beta_4 \cdot DummyTrump_{it} + \beta_5 \cdot TariffRate_{it} + \varepsilon_{it}$$

Share of Income Held By Top 10% (without Tariff Rate):

$$Top10Share_{it} = \alpha + \beta_1 \cdot \log(GDPperCapita_{it}) + \beta_2 \cdot FDI_{it} + \beta_3 \cdot Openness_{it} + \beta_4 \cdot DummyTrump_{it} + \varepsilon_{it}$$

Share of Income Held By Top 10% (with Tariff Rate):

$$Top10Share_{it} = \alpha + \beta_1 \cdot \log(GDPperCapita_{it}) + \beta_2 \cdot FDI_{it} + \beta_3 \cdot Openness_{it} + \beta_4 \cdot DummyTrump_{it} + \beta_5 \cdot TariffRate_{it} + \varepsilon_{it}$$

In this analysis, four separate linear regressions were run in two separate batches. In the first batch, one linear regression was run utilizing the GINI coefficient as a dependent variable while the other linear regression was run using the share of income held by the top 10% as the dependent variable. Both of these linear regressions used the same independent variables of GDP per capita in South Korea, FDI in South Korea, the degree of openness in South Korea, and a dummy Trump variable. In the second batch, while the two different linear regressions were run in the same manner as the first batch in terms of dependent variables, the average tariff rate, which the U.S. imposed on other countries during that period of time, was included in each of the two regressions as an additional independent variable.

The data utilized for both dependent and independent variables were collected for the period between 1980 to 2025. This time frame is particularly significant, as it captures Korea's transition from an authoritarian regime to a democratic country during the 1980s, easing the difficulties that an authoritarian regime may have had on the economy in the form of political instability and decreased economic freedom (Korean Culture and Information Service). The 1980s also marked a period of rapid expansion in trade between the U.S. and Korea, as the U.S. increasingly relied on deficit spending (Nam 189, 201). This allowed us to have a larger volume of data that could be used to analyze trade relations between the two countries. However, some

variables, such as the measures of inequality and the openness indicator, have a couple of years with missing data due to unavailability of complete datasets online.

In order to measure the effect of income inequality in Korea, the GINI coefficient and the share of income held by the top 10% were used as two separate dependent variables. The GINI coefficient is a statistical measure of income inequality, using a range from 0 to 1 in order to assess the income distribution within a country. The GINI coefficient uses 0 to represent perfect equality where everybody in a nation earns the exact same amount and 1 to represent perfect inequality where one sole person has all the income in a country. Meanwhile, the share of income held by the top 10% measures how much of a country's total income is earned by the wealthiest top 10% in the nation. By distinguishing the two measures of inequality and analyzing the outcomes separately, we can assess whether a change in inequality was driven by change within the middle class or a change between the top earners and the rest of the population.

The independent variables for the regressions include GDP per capita yearly in South Korea, FDI yearly in South Korea, the trade openness indicator yearly in South Korea, a dummy Trump variable, and the average tariff rate yearly in the United States. In all four regressions, GDP per capita is converted to its natural log in order to compress the large values into smaller numbers which can increase the stability of the model. Additionally, in all four regressions, the dummy Trump variable utilizes 0 to represent years when a Democrat was president while 1 is used to represent years when a Republican was president. Although tariff policies are not uniform in Democratic or Republican presidencies, and there is no clear distinction in types of tariffs implemented during each party's administration, other economic variables such as tax rates and deficit spending, which have historically been uniform and differentiated between Republicans and Democrats, can be measured through this dummy variable.

Data Analysis

Results

Without Tariff Rates						
	Gini Index			Share of Income Held By Top 10%		
	Co-efficient	Std Deviation	T-Ratio	Co-efficient	Std Deviation	T-Ratio
Const	0.4457***	0.0233	19.10	0.1583***	0.0379	4.179
GDP Per Capita	-0.0190***	0.0029	-6.442	0.0089	0.0052	1.72
FDI	0.0035**	0.0006	5.415	-0.0024**	0.0008	-2.887
Openness	0.0003	0.0002	1.691	0.0023***	0.0002	9.172
Dummy Trump	0.0042	0.0026	1.624	0.0038	0.004	0.943

Table 1: Raw OLS regression results from comparing independent variables listed on the left to two measurements of income inequality (without the introduction of tariffs as an independent variable).

With Tariff Rates						
	Gini Index			Share of Income Held By Top 10%		
	Co-efficient	Std Deviation	T-Ratio	Co-efficient	Std Deviation	T-Ratio
Const	0.8288***	0.1002	8.274	0.5885***	0.1711	3.44
GDP Per Capita	-0.0554***	0.0102	-5.456	-0.025	0.0172	-1.453
FDI	0.001	0.0008	1.143	-0.0042***	0.001	-4.12
Openness	0.00065**	0.0002	2.853	0.00216***	0.00044	4.942
Dummy Trump	0.0032	0.0023	1.376	-0.0013	0.0058	-0.222
Tariff Rate	-0.0177***	0.0051	-3.478	-0.0356**	0.0144	-2.462

Table 2: Raw OLS regression results from comparing independent variables listed on the left to two measurements of income inequality (with the introduction of tariffs as an independent variable).

The four regressions seem to demonstrate strong explanatory power due to their high R-squared values, which range from 0.84 to 0.97. The regression measuring the GINI Index without tariffs as an independent variable has the lowest R-squared value of 0.84; the regression measuring share of income held by the top 10% without using tariffs has an R-squared value of 0.97; the regression measuring the GINI Index using tariffs as an independent variable has an R-squared value of 0.92; the regression measuring the share of income held by the top 10% with tariff rates has an R-squared value of 0.96. As the R-squared value measures how well the independent variables in the regression explain the variation in inequality, with 0 meaning the model explains none of it and 1 meaning that the model explains all the variation, these high values across the regressions show that the independent variables in the models explain the changes in inequality in South Korea. Additionally, both of the regressions that involved the tariff rates had R-squared values of over 90% and noticeably increased the R-squared value of the GINI Index, implying that the incorporation of tariff rates into the models helped better explain the variance in inequality meaning that tariffs in the US are important in explaining the inequality in South Korea.

Summary Table				
Variables	GINI	Top 10%	Gini + Tariff	Top 10% + Tariff
GDPpc	- (***)	+ (not sig)	- (***)	- (not sig)
FDI	+ (***)	- (**)	+ (not sig)	- (***)
Openness	+ (not sig)	+ (***)	+ (**)	+ (***)
DummyTrump	+ (not sig)	+ (not sig)	+ (not sig)	- (not sig)
Tariff Rate	N/A	N/A	- (**)	- (**)

Table 3: Summary of findings from tables one and two including both negative or positive correlation along with levels of significance

In the table, we can observe that GDP per capita is statistically significant at the 1% level and is negative for each of the regressions that used the GINI Index as the dependent variable. This implies that as South Korea has grown economically, overall inequality has decreased, which is consistent with current literature, specifically the Kuznets Curve where per capita GDP increases as overall inequality decreases in developed economies. However, we can also observe that GDP per capita is statistically insignificant to the share of income held by the top 10%, meaning that income concentration among wealthy individuals and top earners is less sensitive to overall economic performance. This means that while overall inequality amongst the middle class decreases through GDP per capita growth, the inequality amongst elites in the top 10% remains unaffected by GDP per capita growth.

Meanwhile, FDI behaves differently in all four regressions. While it was positively significant to the 1% level when just observing the GINI Index, when tariff rates were introduced as an independent variable, FDI became statistically insignificant. This suggests that although foreign capital flow may increase inequality, this effect may be diminished when other international variables such as tariffs are factored in. Meanwhile, FDI remains statistically significant and negatively associated with the share of income held by the top 10% even when tariffs are involved. Thus, although overall inequality may increase or be unaffected by FDI, the income concentration in the hands of the top 10% will be reduced.

Excluding the first regression, which uses the GINI Index as the dependent variable, openness remains positive and statistically significant across the other three regressions, with it being consistently significant to the 1% level for share of income held by the top 10%. This supports the idea in the literature that an increasingly globalized economy can contribute not only to rising income inequality but also to the concentration of wealth amongst the very rich. As stated by Furusawa, Kellar, and Utar, these widening income gaps may result from globalization disproportionately helping those who have large amounts of capital, high-level skills, or employment in companies that are export-oriented. Additionally, openness becomes statistically significant when tariffs are introduced in the regression that uses the GINI Index, suggesting that the effect openness has on inequality can become more pronounced when trade barriers or shocks such as tariffs are accounted for. In other words, this means that globalization and its effects can be significantly pronounced when companies face trade barriers such as tariffs. Thus, in the future, researchers could potentially create a variable that combines the effects of tariffs and openness in order to measure this potentially compounding effect.

The inclusion of the tariff rates for the third and fourth regressions yield the most interesting results. In the regressions where tariffs are included, tariff rates are both negative and statistically significant at the 5% level. This suggests that higher U.S. tariff rates are associated with reductions in South Korean income inequality, a finding that may seem counterintuitive considering that protectionist measures are usually seen as economically harmful. A possible explanation for this is that high tariffs hurt the wealthy the most compared to low-income workers, resulting in reduced income concentration amongst the wealthy even if overall income for the country decreases or stagnates. This interpretation aligns with sources from Lee,

Furusawa, Kellar, and Utar, which show that trade liberalization caused increasing levels of income inequality due to export-oriented large companies benefiting the most from having larger markets and increased competition. Thus, if the opposite of trade liberalization comes in the form of tariffs, the increase in income for the wealthy through the expansion of large export-oriented companies may diminish. Another possible explanation for this result is that higher tariffs, which cause increasing domestic consumption, will force the government or firms to invest in industry diversification, leading to a more balanced distribution of income across multiple sectors. It is also important to note that the effect of globalization and trade has changed significantly from 1980 to 2025. As a result, different economies, such as the South Korean export-focused economy in the 1990s, could cause the inequality-reducing effect to be more pronounced. Overall, these results display how important trade policy is and how it could disrupt not just economic growth, but also many other economic or societal factors such as income inequality.

Finally, the DummyTrump variable remains statistically insignificant for all four regressions. These results suggest that the economic policies that are separated through party affiliation alone, such as the different ideologies on taxes or deficit spending between Democrats and Republicans, are not a good indicator of whether there will be any changes to income inequality in South Korea. This is reasonable as party affiliation only measures general attitudes about policy and not any sort of specific economic instrument. Rather, from a policy perspective, one of the most important factors that affected income inequality in South Korea throughout the regression were tariffs. Thus, although the DummyTrump variable is statistically insignificant, Trump's policy intentions and their effects on South Korean income inequality can be captured in the tariff rate variable rather than the DummyTrump variable which merely differentiates Democratic and Republican presidencies.

Limitations

While the results of this analysis are able to provide some insights into the relationships between potential U.S. Trump tariffs on South Korean income inequality, there are some limitations in this paper that could be improved upon in the future. One major limitation is the amount of data that could be collected online. Although the use of quarterly or monthly periods could help reveal shorter term fluctuations, lagged effects for variables that do not display their effects immediately, or the effect that short-term tariffs may have, there was a lack of data with these short time periods. Additionally, because there are currently only four observable years during which Donald Trump has been in office, there were not enough data points to focus the DummyTrump variable on his presidency, causing the variable to be relatively simplistic. In the future, when Trump serves 8 years in office, it would be valuable to revisit this topic with more data points to explore the specific effects of his presidency, rather than just his tariff policies, on South Korean inequality. Finally, as there was not any data online displaying the tariffs that the U.S. has had on Korea specifically from 1980 to 2025, I was forced to use U.S. overall tariffs despite there being a possible gap between Korea specific and U.S. overall tariff rates. Thus, the

overall lack of data may limit the ability to explore many promising measures related to this topic or introduce minor measurement errors into the regression analysis.

The OLS method also assumes that the relationship between each variable and inequality is a straight line. However, there might be a nonlinear relationship between inequality and these independent variables. Thus, future methods could explore if other models or methods could capture a more complex dynamic between each of the independent variables and inequality.

Conclusion

The analysis of the relationship between Donald Trump's potential tariff policies and income inequality in South Korea indicates that an increase in U.S. tariffs will cause a significant decrease in income inequality in South Korea, decreasing the gap between the top 10% and the low-wealth class in South Korea. Beyond the relationship between tariffs and income inequality, we found results consistent with the Kuznets Curve where GDP per capita decreases income inequality for developed countries. However, these results showed that while overall GINI was impacted, the change in income gained by the top 10% remained insignificant, translating to a decrease in income inequality mainly amongst the middle class. The results of FDI on inequality were a little inconsistent for the GINI Index, but we found that FDI consistently decreased inequality amongst the top 10%, even when variables like tariffs were introduced. We also found that an increase in openness nearly consistently led to an increase in income inequality, especially so when trade barriers such as tariffs were introduced. This suggests that the effect that openness has on income inequality may be even more pronounced when trade barriers are accounted for. Finally, we found that the Dummy Trump variable which measured Republican presidencies and broad Republican tax trends remained consistently insignificant for all measures of inequality, highlighting the idea that it is tariffs rather than Republican economic policy that mainly affect the inequality in South Korea. However, it is extremely important to note that the results gleaned from analyzing the interaction between U.S. tariffs and South Korean inequality do not mean tariffs should be utilized as a tool to decrease income inequality. Rather, the use of protectionist measures can significantly disrupt overall economic growth for both of the countries involved. These findings show that foreign trade policy is not just important to economic growth but can have large implications for other societal or economic issues, such as the crucial societal issue of income inequality. As global tensions continuously evolve over tariffs, it is essential that policymakers worldwide understand the dangers, benefits, and importance that tariffs have on economic and social issues, beyond mere economic growth.

Future research could potentially explore the use of sector-specific tariff data, utilize Korean domestic economic policies, or postpone analysis until both of Trump's terms have concluded to fully capture the drivers of inequality and their relation to tariffs.

Works Cited

- Lee, Siwook. "International Trade and Within-Sector Wage Inequality: The Case of South Korea." *Journal of Asian Economics*, vol. 48, Feb. 2017, pp. 38–47. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.asieco.2016.11.001>.
-
- Cheong, Juyoung, and SeEun Jung. "Trade Liberalization and Wage Inequality: Evidence from Korea." *Journal of Asian Economics*, vol. 72, Feb. 2021, p. 101264. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.asieco.2020.101264>.
- Savvides, Andreas. "Trade Policy and Income Inequality: New Evidence." *Economics Letters*, vol. 61, no. 3, Dec. 1998, pp. 365–72. *DOI.org (Crossref)*, [https://doi.org/10.1016/S0165-1765\(98\)00197-9](https://doi.org/10.1016/S0165-1765(98)00197-9).
- Jung, Hyun-Uk, and Tae-Hyoung Mun. "The Effect of U.S. Protectionist Trade Policy on Foreign Ownership: A Study of Korea's Data Set." *Journal of Korea Trade*, vol. 23, no. 7, Nov. 2019, pp. 83–95. *DOI.org (Crossref)*, <https://doi.org/10.35611/jkt.2019.23.7.83>.
- Wang, Meiling, et al. "The Nexus between International Trade, FDI and Income Inequality." *Journal of Korea Trade*, vol. 24, no. 4, June 2020, pp. 18–33. *DOI.org (Crossref)*, <https://doi.org/10.35611/jkt.2020.24.4.18>.
- Rosenfeld, Tomas, et al. "International Trade and Income Inequality: The Case of Latin American Countries." *The International Trade Journal*, vol. 38, no. 5, Sept. 2024, pp. 469–89. *DOI.org (Crossref)*, <https://doi.org/10.1080/08853908.2024.2341624>.
- Aradhyula, Satheesh V., et al. *Impact of International Trade on Income and Income Inequality*. 2007. *DOI.org (Datacite)*, <https://doi.org/10.22004/AG.ECON.9999>.
- Furusawa, Taiji, et al. "International Trade and Income Inequality*." *The Scandinavian Journal of Economics*, vol. 122, no. 3, July 2020, pp. 993–1026. *DOI.org (Crossref)*, <https://doi.org/10.1111/sjoe.12360>.
- Keller, Wolfgang, and Hale Utar. "International Trade and Job Polarization: Evidence at the Worker Level." *Journal of International Economics*, vol. 145, Nov. 2023, p. 103810. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.jinteco.2023.103810>.
- Shin, Kwang-Yeong, and Ju Kong. "Why Does Inequality in South Korea Continue to Rise?" *Korean Journal of Sociology*, vol. 48, no. 6, Dec. 2014, p. 31. *DOI.org (Crossref)*, <https://doi.org/10.21562/kjs.2014.12.48.6.31>.
- Hong, Minki. "Structural Changes in South Korean Employment (2000–2021)." *Global Trends in Job Polarisation and Upgrading*, edited by Sergio Torrejón Pérez et al., Springer Nature Switzerland, 2025, pp. 297–320. *DOI.org (Crossref)*, https://doi.org/10.1007/978-3-031-76228-4_11.
- Yoon, L. "South Korea: Social Welfare Expenditure." *Statista*, 24 Aug. 2024, <https://www.statista.com/statistics/618089/south-korea-social-welfare-expenditure/>.
- Wile, Rob, et al. "Trump Threatens 200% Tariff on E.U. Wine and Liquor, Says World Is 'Ripping Us Off.'" *NBC News*, 13 Mar. 2025,

- <https://www.nbcnews.com/business/economy/trump-threatens-200-tariff-eu-booze-tax-u-s-whiskey-rcna196192>.
- Service (KOCIS), Korean Culture and Information. *Transition to a Democracy and Transformation into an Economic Powerhouse : Korea.Net : The Official Website of the Republic of Korea*.
<https://www.korea.net/AboutKorea/History/Transition-to-a-Democracy>. Accessed 29 Mar. 2025.
- Nam, Chong-Hyun. "Protectionist U.S. Trade Policy and Korean Exports." *Trade and Protectionism*, University of Chicago Press, 1993, pp. 183–222. *National Bureau of Economic Research*,
<https://www.nber.org/books-and-chapters/trade-and-protectionism/protectionist-us-trade-policy-and-korean-exports>.
-
- "GDP per Capita, Current Prices (South Korea)." *IMF*, 2024,
www.imf.org/external/datamapper/NGDPDPC@WEO/KOR?zoom=KOR&highlight=KOR.
-
- South Korea Foreign Direct Investment 1970-2025*.
<https://www.macrotrends.net/global-metrics/countries/kor/south-korea/foreign-direct-investment>. Accessed 29 Mar. 2025.
- Korea Gini's Coefficient Index: All Households: Disposable Income | Economic Indicators | CEIC*.
<https://www.ceicdata.com/en/korea/household-income-and-expenditure-survey-hs-income-distribution/ginis-coefficient-index-all-households-disposable-income>. Accessed 29 Mar. 2025.
- "World Bank Open Data." *World Bank Open Data*,
<https://data.worldbank.org/indicator/SI.POV.GINI?locations=KR>. Accessed 29 Mar. 2025.
- kyungsoo CHOI. Explaining Income Inequality in Korea July 2003. 2003.
- "Share of Income Held By Top 10% Data." *WID - World Inequality Database*,
https://wid.world/data/#countrytimeseries/sfiinc_p90p100_z;sfiinc_p99p100_z/KR/1994/2016/eu/k/p/yearly/s. Accessed 29 Mar. 2025.
- Openness at Constant Prices for Republic of Korea*. 17 Sept. 2012,
<https://fred.stlouisfed.org/series/OPENRPKRA156NUPN>.
- U.S. Tariff Rates 1989-2025*.
<https://www.macrotrends.net/global-metrics/countries/usa/united-states/tariff-rates>. Accessed 29 Mar. 2025.
-

Sustainable System for Removing Toxic Skid Marks from F1 Racing on Singapore Streets through Waste Upcycling By Benjamin Frederick Oh

Abstract

As Formula 1 (F1) gains popularity, its sustainability issues are increasingly coming under scrutiny. While various environmental concerns remain detrimental, tire skid marks are one of particular concern for Singapore due to its racing circuit's proximity to the sea. There is a potential risk that harmful chemicals from these skid marks could enter the marine ecosystem. This study explored an eco-friendly approach to removing skid marks left on public roads during the Formula 1 Grand Prix in Singapore by upcycling local waste materials. The Skid Mark Scratcher (SMS) was developed using various waste materials such as coconut fiber, cigarette filter, clam shells, and gelatin. The SMS effectively broke down skid marks on, and the fragments were easily vacuumed. Additionally, bacteria such as *Rhodococcus* sp. and soil bacteria were tested for their ability to decompose the toxins of skid mark chemicals like dimethyl disulfide and zinc oxide, which are harmful to human neuroblastoma cells and lung cells. In conclusion, the experiment confirmed that SMS is effective for removing skid mark and showed that these functional bacteria could mitigate the harmful effects of chemicals. In real-life applications, this sustainable system, integrating SMS and microbial agents, suggests the potential for further improvement in effectiveness.

1. Introduction

The popularity of Formula 1 (F1) racing has surged globally, with Singapore's Grand Prix standing out as one of the most prestigious and exciting events on the F1 calendar. This is evidenced by the 2021 season drawing a cumulative TV audience of 1.55 billion and F1's social media followers increasing by 40 % to 49.1 million [1]. The Singapore Grand Prix, unique for its night race and urban track layout, attracts fans and participants worldwide, significantly boosting tourism and the local economy by "drawing... teams" and being "attractive to visitors" [2]. However, the environmental impact of F1 races, including the significant rubber deposits left on the tracks, fuel consumption, and emissions, cannot be overlooked. In 2023, F1 left a carbon footprint of 223,031 tons (CO₂), highlighting the sport's substantial environmental challenges [3].

After each race, rubbers from tire skids are left on the racetracks, and these skid marks are composed of complex chemicals such as sulfur compounds, which are used for sulfur vulcanization, and zinc oxide [4]. These chemicals pose serious threats to environment and human. When these materials degrade, their byproducts can leach into the soil and waterways, leading to long-term ecological consequences. On roads, these chemicals can be washed off during rainfall, contaminating nearby water sources, which can harm aquatic life and disrupt ecosystems. A study on nano-ZnOs, one of the chemical compounds in tires, has shown that high-level exposure to ZnO can significantly impact organ function and mineral biodistribution in humans and animals [5].

Similarly, dimethyl disulfide (DMDS), the other chemical compound in tires, is also known for being highly flammable, toxic, and very harmful to aquatic life [6]. Therefore, prolonged exposure to these contaminants can also affect the health of animals that encounter polluted water or soil, potentially causing detrimental effects on the ecosystem. Especially in Singapore, it is essential to apply a method to mitigate the toxicity of road pollutants before they leak into the sea because the proximity of the racetrack to the sea exacerbates the risk of marine contamination. While the F1 industry has made strides in addressing carbon emissions and other environmental issues, tackling the rubber residue problem requires innovative solutions.

In Singapore, cigarette butts, coconut husks, and straw waste pose significant environmental challenges due to their high accumulation rates and persistence in local ecosystems. Cigarette butts are one of the most frequently collected litter items along Singapore's coastlines, accounting for around 20,915 pieces during a single cleanup event, nearly double the count for plastic straws [7]. These pollutants further complicate the environmental landscape and offer an opportunity to develop a comprehensive approach to cleaning racetracks. By repurposing marine waste in physically removing skid marks, we can create a more sustainable method of managing rubber residues.

In addition to physical removal methods, biological approaches offer promising solutions for mitigating the environmental impact of skid marks that could significantly reduce the ecological footprint of F1 races. *Rhodococcus* sp. is utilized for the bioremediation of ZnO and DMDS toxicity, capitalizing on the bacteria's unique capabilities in both metal and sulfur compound degradation [6]. *Rhodococcus* is a resilient, aerobic, Gram-positive bacterium known for its adaptability in diverse and often harsh environmental conditions, making it particularly effective in treating soil and water contaminants [8]. Studies have demonstrated that *Rhodococcus* can transform heavy metals like ZnO, accumulating and neutralizing their toxic effects in soil and water, thus reducing bioavailability and potential harm to ecosystems [8]. Additionally, *Rhodococcus* is specialized in breaking down sulfur compounds such as DMDS through oxidative pathways, which convert these substances into less toxic forms, thereby addressing environmental sulfur pollution [9]. This dual capability of *Rhodococcus* in transforming both ZnO and sulfur compounds supports its application in pollution control, as it can simultaneously target multiple toxins, ensuring broad-spectrum bioremediation efficiency in complex polluted environments.

Therefore, this research aims to explore physical and biological methods to effectively remove tire remnants and skid marks from racetracks, specifically focusing on the Singapore Grand Prix. By integrating innovative physical removal techniques with biological degradation processes, this study seeks to develop a sustainable solution that addresses the environmental impact of rubber residues and contributes to reducing pollution in terrestrial and marine ecosystems.

2. Materials and Methods

2.1. Skid Mark Scratcher (SMS)

2.1.1. Prepare the Materials for SMS

The extraction of gelatin from pork skin (Figure 1-A) was carried out. Initially, 200 g of pork grind was cleaned and cut into small pieces (upper panel in Fig 1-B). These pieces were slightly boiled for 5 minutes to remove surface impurities. After boiling, the pork grind was rinsed thoroughly under running water. The cleaned pork grind was then boiled in a solution of water and vinegar at 160 °C for 3 hours (lower panel in Fig 1-B). Throughout this boiling process, water was periodically added to maintain the volume. Once the boiling was complete, the mixture was filtered to extract the gelatin (Figure 1-C), which was then stored in the refrigerator for future use.

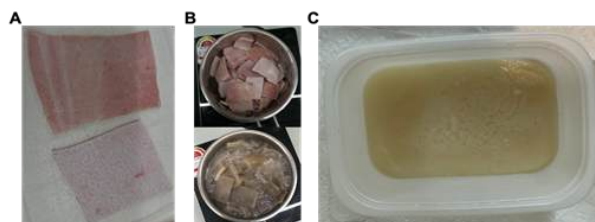


Figure 1. Isolation of gelatin from pig skin (A, pork skin; B, boiling pork skin cut; C, gelatin extracted from pork skin)

Clam shells purchased from an online shopping mall were thoroughly washed to remove any impurities. Once cleaned, the clam shells were dried completely at room temperature in a fume hood. The dried clam shells were then ground into powder using a blender. Sawdust was completely dried to remove any moisture content. Raw sawdust was spread out evenly on a tray and placed in an oven set to 50 °C. The drying process was carefully monitored to ensure the sawdust was thoroughly dried without burning. Once dried, the sawdust was stored in a clean, dry container.



Figure 2. Clearance of cigarette butt filters (A, cigarette butt in water; B, cigarette butts with water removed; C, cigarette butt decomposition; D, separation of cigarette butt filters; E, filters in 70% ethanol; F, completely dried cigarette butt filters)

The collected cigarette butts were soaked in water to remove residual chemicals and soften the material (Figure 2A & B). After sufficient soaking, cigarette butt filters were easily separated from the butts (Figure 2-C & D). The filters were soaked in 70% ethanol to remove the smell (Figure 2F), and then filters were drained and dried (Figure 2-E). The dried filters were

then stored in a container. Coconut fibers were collected and cut into shorter lengths. The shorter coconut fibers were then dried and stored in a clean container for use.

2.1.2. Construction of SMS

5 mL each of ground clam shell, sand and sawdust were transferred to a 50 mL conical tube (Figure 3). The coagulated gelatin that had been stored in the refrigerator was melted by heating it in a microwave oven and then 15 mL of melted gelatin was added to the tube. The mixture was mixed well by vortexing.

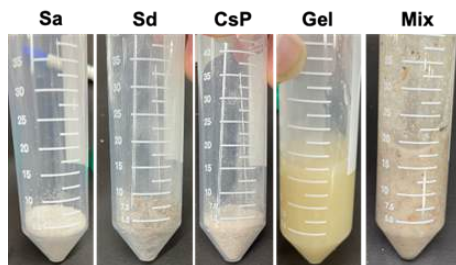


Figure 3. Preparation of skid mark scratcher mixture (Sa = sand; Sd = sawdust; CsP = clam shell powder; Gel = gelatin; Mix = mixture)

Three 6-well plates were prepared with different bases: the first plate was evenly distributed with dried cigarette butt filters (Cr), the second with shorter coconut fibers for easier distribution (Cf), and the third combined both cigarette butts and coconut fibers (Cr+Cf) (Figure 4-A). The prepared mixture was poured into each well (Figure 4-B) and allowed to dry, forming the tire skid scratchers.

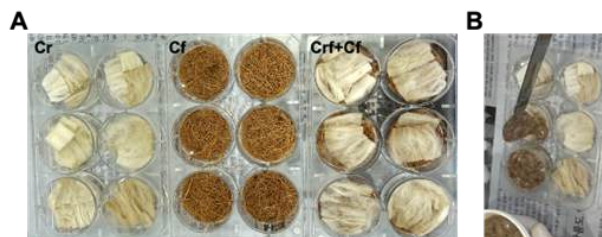


Figure 4. Construction of skid mark scratcher (A, cigarette butt filters, coconut fibers, or cigarette butt filters and coconut fibers in 6-well plate; B, addition of mixture of sand, sawdust, clam shell powder and gelatin)

2.2. Simulation of Skid Mark on the Road

2.2.1. Cement-Based Road

The initial step for simulating the racing road was to construct the cement bases. 150 mL each of cement, sand, and water were combined in a mixing container. The mixture was thoroughly stirred using a stirring rod until a uniform consistency was achieved. The prepared cement mixture was divided equally into three portions and poured into separate Petri dishes, ensuring even distribution across the bottom of each dish. To create cement-crushed stone bases, approximately 33 mL of crushed stone was added to each dish immediately after pouring the

cement mixture. The mixture was gently pressed to embed the crushed stone, forming a flat, uniform layer. The cement and cement crushed stone bases were allowed to set and cure at room temperature (Figure 5-A).

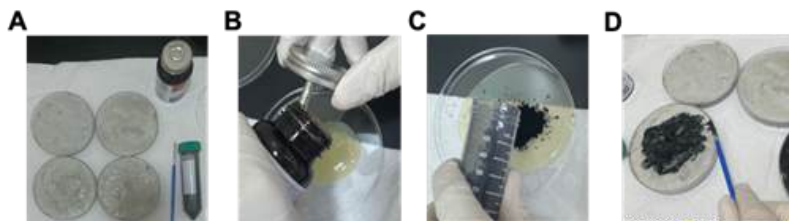


Figure 5. Skid mark simulation (A, prepared cement-based roads; B, addition of rubber cement in Petri dish; C, addition of charcoal powder; D, applying the mixture of rubber cement and charcoal powder on the cement-based road)

2.2.2. Skid Mark Simulation

Rubber cement and charcoal powder were purchased from the online shopping mall (COUPANG). The rubber cement and charcoal were mixed in a Petri dish (Fig 5-B & C). The mixture was applied using a brush to create various skid mark patterns on both the cement and cement-crushed stone bases (Fig 5-D). Three different types of skid marks were created: thin skid marks, thick skid marks, and circular skid marks covering the entire Petri dish. These models were left to dry, providing realistic skid mark simulations for further analysis.

2.3. Skid Mark Scratcher Testing

The skid marks simulated on a concrete-based road were scratched using a prepared SMS (Figure 6-A). A small vacuum cleaner was used to recover fragments that were detached from the skid marks and SMS (Figure 6-B). Following the scratching, a mini vacuum cleaner was used to suck the detached fragments of skid marks and scratcher (Figure 6-C).

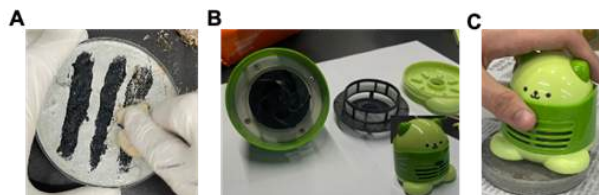


Figure 6. Removal of skid marks on the cement-based road by the SMS and suction of skid mark fragments by a vacuum cleaner (A, scratching the skid mark by using the SMS; B, a small vacuum cleaner; C, recovering the fragments of skid mark and SMS)

2.4. Cultivation of *Rhodococcus* sp. or soil bacteria

Rhodococcus sp. (KCTC 19690) was purchased from Korea Collection for Type Cultures (KCTC). Using a glass cutter, the glass container with freeze-dried *Rhodococcus* sp. was carefully opened (Figure 7-A). The bacteria were suspended with 100 μ L of nutrient broth (NB)

media (Figure 7-B). 10 mL of NB media was transferred to a new 50 mL conical tube (Figure 7-C), and the resuspended bacteria was added to the tube (Figure 7-D).

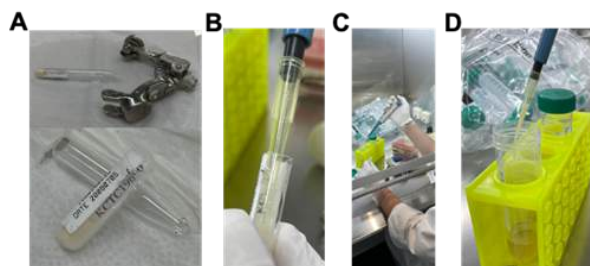


Figure 7. Cultivation of *Rhodococcus* sp. (A, cutting the glass container with *Rhodococcus* sp. by using a glass cutter; B, addition of NB media into the cut glass container; C, transferring the suspended *Rhodococcus* sp. to a new 50 mL conical tube; D, inoculation of the bacteria solution to the tube)

20 mL of soil from the flower bed in front of the house was transferred to a 50 mL tube containing 20 mL of distilled water (DW). After mixing by inverting at room temperature for 1 hour, 1 mL of supernatant was transferred to each 50 mL tube filled with 10 mL of tryptic soy broth (TSB) media. The growth of soil bacteria was checked by observing the change in turbidity of broth media.

2.5. Preparation of ZnO and dimethyl disulfide Solutions

ZnO or dimethyl disulfide (DMDS) were purchased from Ducksan General Science or Sigma-Aldrich, respectively. To prepare the 100 mM ZnO stock solution, 0.41 g of ZnO was dissolved in 1 mL of 35% HCl to ensure complete dissolution. Afterward, DW was added to adjust the final volume to 50 mL. For the negative control in this study, a 0.7% HCl solution was used, corresponding to the acid concentration in the final 100 mM ZnO solution. 100 mM ZnO or 100% DMDS was diluted to various concentrations by Dulbecco's Modified Eagle's Medium containing antibiotics (penicillin and streptomycin) and 10% fetal bovine serum (FBS) (DMEM (+/+)), or TSB media. To obtain the HCl, ZnO or DMDS solution treated with *Rhodococcus* sp. or soil bacteria, 500 μ L of the precultivated bacteria was transferred to a new 15 mL conical tube with 10 mL of NB or TSB media containing 15 mM ZnO or 1.6% DMDS. The bacteria were incubated at room temperature for 5 days. After centrifuging at 3,000 rpm for 20 minutes, the supernatant was transferred to a new 15 mL tube.

2.6. A549 or SK-N-MC cells culture

2.6.1. Cell Subculture

In this study, A549 cells, human lung adenocarcinoma cell line, and SK-N-MC cells, human neuroblastoma cell line, were used. To subculture A549 cells or SK-N-MC cells, the old medium was removed from the pre-cultivated cells, and the cells were washed with 5 mL of 1x phosphate-buffered saline (PBS). After discarding the PBS, 700 μ L of 0.25% trypsin/EDTA was added to cells. The cells were incubated at 37°C for 3 minutes, and then 1 mL of 1x PBS was

added to cells. The cells were suspended and transferred to a new 15 mL conical tube. After centrifuging at 1,500 rpm for 3 minutes, the PBS-trypsin mixture was discarded, and the cell pellet was resuspended with fresh 1 mL of DMEM (+/+) media. The resuspended cells were diluted with 14 mL of DMEM (+/+) media. Finally, 100 μ L of the cells were added into each well of 96 well plate and incubated at 37 °C in 5%

CO₂. Regular observation of the A549 and SK-N-MC cells was conducted to monitor cell health, adherence, and morphology. The cells in 96-well plates were examined under an inverted microscope at 40x and 100x magnifications.

2.6.2. Treatment the Cells with DMDS or ZnO

The DMEM from the cultivated A549 or SK-N-MC cells was removed, and the cells were washed with 200 μ L of PBS. 200 μ L of the prepared DMEM (+/+) media containing each concentration of ZnO, DMDS or HCl solution was added to A549 or SK-N-MC cells. To compare the reduction in toxicity of ZnO or DMDS by bacteria, 200 μ L of HCl, ZnO or DMDS solution treated with *Rhodococcus* sp. or soil bacteria was added to A549 or SK-N-MC cells, and then the cells were incubated at 37 °C in 5% CO₂ for 2 days.

2.7. Cell Viability Test

The cell viability was determined by using Cell Counting Kit-8 (CCK-8) (Dojindo). A 10 % CCK solution was prepared by mixing 7.2 mL of DMEM (+/+) with 800 μ L of CCK-8 reagent in a 15 mL conical tube. The SK-N-MC or A549 cells treated with each chemical were first washed with 200 μ L of 1x PBS, and then 100 μ L of the prepared 10 % CCK-8 solution was added to cells. After incubating for 15 minutes at 37°C, the value of optical density at 450 nm (OD₄₅₀) of the cells was measured by using a microplate reader.

3. Results

3.1. Effectiveness of Skid Mark Scratcher

In an effort to seek eco-friendly ways to remove skid marks, SMSs were constructed using various natural and recyclable materials, such as cigarette butt filters, coconut fibers, sawdust, and clam shells. The goal of this experiment was to evaluate the effectiveness and durability of these materials in practical applications. Three types of SMS were prepared using different base materials as follows: cigarette butt filters (Crf), coconut fiber (cf), and Crf+Cf (lower panel in Fig 8). It was observed that the Crf- or Cf-based SMS maintained their shape well. However, the Crf and Cf-based SMS showed some signs of demolition, indicating structural weakness when these two materials were combined (Dotted box of Figure 8). Therefore, it was decided that Crf- or Cf-based SMS was used in subsequent simulations.

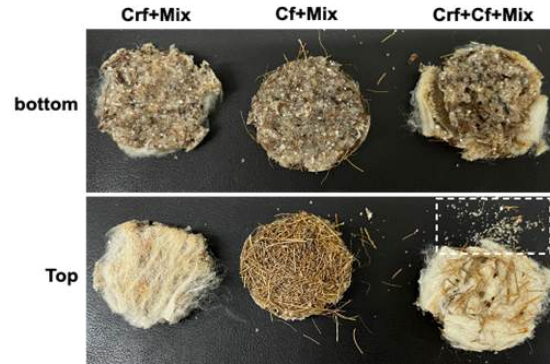


Figure 8. Skid mark scratcher types (Crf, cigarette butt filter; Cf, coconut fiber; Mix, saw dust, sand, clam shell powder and gelatin)

3.2. Adsorption of Skid Mark's Small Debris into Sawdust

To confirm whether the SMSs absorb small debris of skid marks, the simulated skid marks on the cement-based road were scratched by using each SMS, and then the particles from SMS and skid marks were observed under microscope. Three types of skid marks were recreated for this test (Figure 9-A). Each type of SMS was used to scratch across these skid marks, effectively removing the rubber deposits from the surfaces (shown as "a" in Figure 9-B). After this process, the particles detached from the SMS and skid mark were collected (see "b" & "c" in Figure 9-B) and gathered in a 6well plate to observe. (Figure 9-C). This setup enabled an easy and accurate analysis of SMS effectiveness in removing and retaining rubber debris.

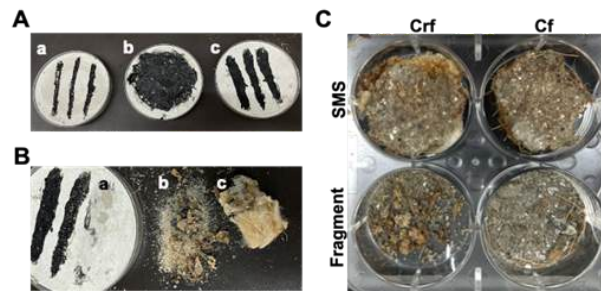


Figure 9. Skid mark residue after application of skid mark scratcher (A, skid marks simulated on cement-based road; B, skid mark after scratching with SMS; C, the base of the SMS remaining after scratching and SMS fragments; a in A = thin skid mark; b in A = the thickest skid mark; c in A = medium thickness skid mark; a in B = scratched skid mark; b in B = skid mark remnants and SMS fragments; c in B = the base of the SMS remaining after scratching; Crf = cigarette butt filter-based SMS; Cf = coconut fiber-based SMS)

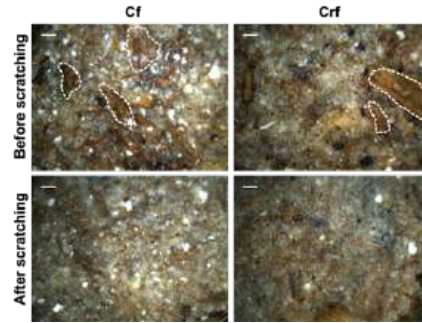


Figure 10. Surface of Skid mark scratcher after removing skid marks (Cf = coconut fiber-based SMS; Crf = cigarette butt filter-based SMS)

The largest particles in SMS, sawdust, were found to largely detach from the surface after friction with the skid marks (Dotted circle of Figure 10). This detachment indicates that sawdust particles were removed from the SMS surface, likely collecting rubber debris as they separated. It suggests a strong possibility that sawdust particles can adhere to and become integrated with the skid mark material during the scratching process. Sawdust is composed of lignin, and it is known that lignin can adsorb a variety of substances, including organic, metal, and gaseous pollutants, dyes, and proteins [10]. Therefore, in this study, sawdust was used to absorb skid mark fragments. The purpose of this experiment was to observe how well skid mark remnants are stuck to sawdust of SMS, as well as how easily stuck remnants are removed. Following the scratching of skid marks on cement-based road using a Crf- or Cf-based SMS, the remnants were observed under microscope. Compared to the negative controls (N.C in Figure 11), which showed no such attachment, the experimental samples (Cf or Crf) clearly demonstrated the effectiveness of both scratchers in attaching skid mark fragments (red arrowhead in Figure 11). This adherence indicates that the skid mark fragments are easier to clean, confirming that both types of scratchers are effective for skid mark removal.

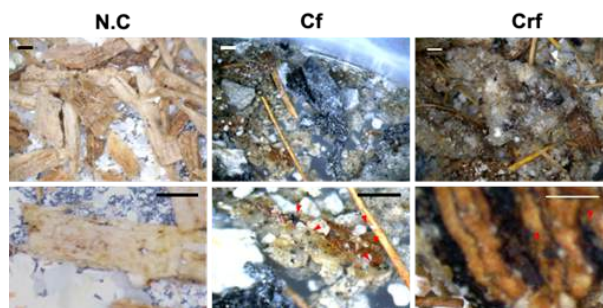


Figure 11. Absorption of skid mark fragment on surface of saw dust (N.C = negative control, a simple mixture of sawdust, seashells and sand; Cf = coconut fiber + simple mixture; Crf, cigarette butt filter + simple mixture; red arrowheads = skid mark fragments attached on the sawdust)

3.3. Suction of Skid Mark Fragments

Since the skid mark fragments and SMS remnants must be promptly removed from the road on which the race had taken place, it was evaluated whether the fragments and remnants can

be effectively cleaned or collected using a vacuum cleaner. Skid marks on the cement-based road were effectively removed with five round trips of the Cf-based SMS (Figure 12-A). A small vacuum cleaner was operated on the cement based road with skid marks removed, for 30 seconds. Almost skid mark fragments (white arrow in Figure 12-B) and SMS remnants (white dashed line in in Figure 12-B) were quickly collected with the small vacuum cleaner. It suggests that operating the SMS in combination with a vacuum cleaner is an effective way to quickly remove skid marks on racetracks.

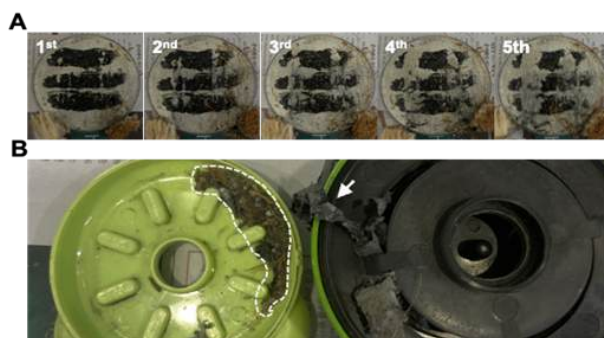


Figure 12. Removal of skid mark and suction of the fragments (A, scratching the skid marks simulated on the cement-based road by SMS; B, collecting the skid mark fragments and the SMS remnants by vacuum cleaner)

3.4. Reduction of Cytotoxicity of ZnO or DMDS by *Rhodococcus* sp. or soil bacteria

3.4.1. Growth of A549 or SK-N-MC cells

To investigate the toxicity reduction capabilities of *Rhodococcus* sp. or soil bacteria on contaminants found in skid marks, human cell lines were used. *Rhodococcus* sp., originated from soil, has been shown to effectively degrade various toxic substances [6]. Skid marks from roads may contain compounds like ZnO or DMDS, both known to exhibit toxicity toward skin or lung cells [5]. Despite physical removal of skid marks, ZnO or DMDS residues can remain on the road and pose potential health risks. Therefore, A549, human lung adenocarcinoma, and SK-N-MC, human neuroblastoma, cell lines were decided to assess whether *Rhodococcus* sp. or soil bacteria alleviated the toxicity of ZnO or DMDS to human lung or nerve cells. For this end, first, A549 or SK-N-MC cells were cultivated.

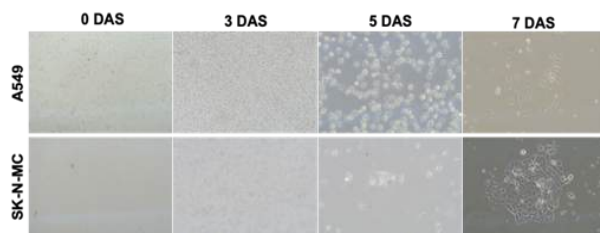


Figure 13. Proliferation of A549 or SK-N-MC cells (0 and 3 DAS, 20x magnification; 5 DAS, 40x magnification; 7 DAS, 200x magnification; DAS = days after seeding)

The results of cultivating A549 or SK-N-MC cell lines over 0, 3, 5, and 7 days after seeding (DAS) demonstrated changes in cell growth and morphology (Figure 13). Over time, both cell lines showed an increase in cell count, which indicates healthy proliferation as the culture period progressed. On Day 0 (initial seeding), cells appeared sparsely distributed with limited attachment. By Day 3, there was noticeable cell proliferation, with cells covering a larger area of the culture dish and beginning to form more compact clusters. By Day 5, cell density further increased, with more cells establishing intercellular connections, particularly in A549 cells, which displayed a more epithelial-like, cobblestone morphology. SK-N-MC cells, meanwhile, adopted a more elongated, spindle-like shape. By Day 7, both cell lines completely adhered to culture flask, with A549 cells maintaining a uniform appearance and SK-N-MC cells continuing to exhibit elongated shapes.

3.4.2. Determination of the Inhibitory Concentration (IC_{50}) of ZnO or DMDS solution

To avoid the inability to determine the cytotoxicity reduction of ZnO or DMDS by *Rhodococcus* sp. or soil bacteria due to the use of too high concentrations of the chemicals, it was required that determine the concentration of ZnO or DMDS which reduces viability of A549 or SK-NMC cells to 50 % (IC_{50}). ZnO was tested on SK-N-MC cells due to its known neurotoxic effects, which include triggering autophagic cell death pathways [11]. In contrast, DMDS, which is associated with lung damage, was tested on A549 cells, which are a lung cell model frequently used to assess respiratory toxicity [12]. By selecting relevant cell lines, the experiment aimed to determine effective concentrations that could model each compound's specific toxicological effects and verify the detoxification potential of *Rhodococcus* and soil bacteria.

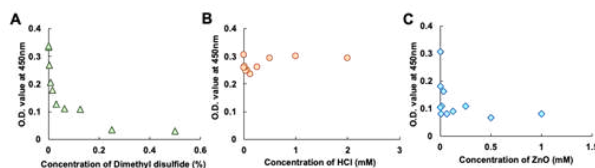


Figure 14. Inhibitory concentration (IC_{50}) of ZnO or DMDS (A, IC_{50} of DMDS in A549 cells; B, IC_{50} of HCl in SK-N-MC cells; C, IC_{50} of ZnO in SK-N-MC cells)

The results demonstrated varying levels of cytotoxicity among the tested chemicals. For DMDS, the IC_{50} for A549 cells was determined to be 0.016 %, indicating 0.016 % of DMDS reduces the survival of A549 cells to 50%. (Figure 14-A). In contrast, each diluted HCl solution showed no significant cytotoxicity against SK-N-MC cells, with consistent OD_{450} values observed across all tested concentrations, confirming that the concentration of HCl for dissolving ZnO did not negatively affect cell viability (Figure 14-B). For ZnO, the concentration of 0.15 mM was identified as IC_{50} for SK-N-MC cells, indicating that 0.15 mM of ZnO reduces the survival of SK-N-MC cells to 50% (Figure 14-C). These findings highlight the specific concentrations at which DMDS or ZnO can be used in future testing without inducing excessive cytotoxicity, while also confirming the safety of the HCl solution at the tested dilutions.

3.4.3. Effect of Bacteria on Reducing Cytotoxicity of ZnO or DMDS

The purpose of this experiment was to assess whether *Rhodococcus* sp. or soil bacteria reduce the toxicity on A549 or SK-N-MC cells of ZnO or DMDS. ZnO or DMDS solution was incubated with *Rhodococcus* sp. or soil bacteria, and then A549 or SK-N-MC cells were treated with each solution. ZnO or DMDS solution diluted with NB or TSB media was used for negative control.

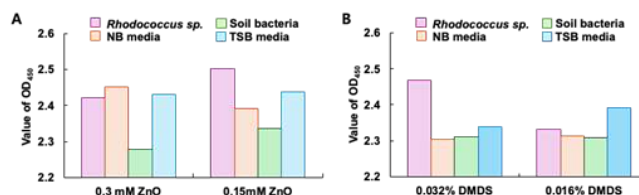


Figure 15. Reduction of toxicity of ZnO (A) or DMDS (B) on A549 cells by *Rhodococcus* sp. or soil bacteria

In A549 cells, the OD₄₅₀ value of cells incubated with 0.3 mM ZnO treated with *Rhodococcus* sp. did not significantly differ to the value of cells incubated with only 0.3 mM ZnO diluted with NB media (Figure 15-A). The value of cells incubated with 0.3 mM ZnO treated with soil bacteria was lower than the value of cells treated with 0.3 mM ZnO diluted with TSB media (Figure 15-A). Interestingly, the value of cells incubated with 0.15 mM ZnO treated with *Rhodococcus* sp. was higher than those of the other samples (Figure 15-A). It indicates that *Rhodococcus* sp. is effective in reducing the toxicity of ZnO to A549 cells. Regarding DMDS exposure in A549 cells (Figure 15-B), at 0.032% DMDS, *Rhodococcus* sp. treatment resulted in a high OD₄₅₀ value, indicating a reduction in DMDS toxicity, whereas the other treatments (soil bacteria, NB, and TSB) showed moderate values. At 0.016% DMDS, TSB yielded a slightly higher OD₄₅₀ value than the other treatments, which showed moderate values. It may suggest that in high concentrations of DMDS, *Rhodococcus* sp. is effective, but in lower concentrations, none of the treatments are effective.

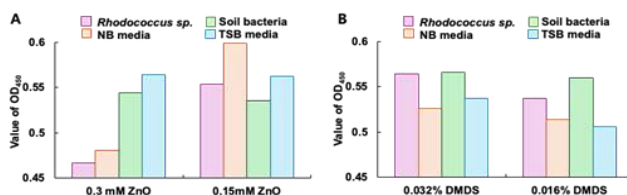


Figure 16. Reduction of toxicity of ZnO (A) or DMDS (B) on SK-N-MC cells by *Rhodococcus* sp. or soil bacteria

The OD₄₅₀ value of SK-N-MC cells treated with 0.3 mM ZnO, whether cultivated with or without *Rhodococcus* sp., exhibited consistently low (Figure 16-A). The value of SKN-MC cells treated with soil bacteria-treated 0.3 mM ZnO was slightly lower than those treated with ZnO diluted in TSB media. For 0.15 mM ZnO, the value of cells treated with *Rhodococcus* sp.-treated ZnO was slightly lower than the value of cells treated with ZnO diluted with NB media. Additionally, the value of SK-N-MC cells treated with soil bacteria-treated 0.15 mM ZnO was slightly lower than the value of cells treated with ZnO diluted with TSB media. These data

suggest that *Rhodococcus* sp. or soil bacteria are not suitable for reducing the cytotoxicity of ZnO to SK-N-MC cells. The OD₄₅₀ values of SK-N-MC cells incubated with 0.032% or 0.016% DMDS treated with *Rhodococcus* sp. or soil bacteria were higher than the those of cells treated with DMDS diluted in NB or TSB medium. These results collectively indicate that both *Rhodococcus* sp. and soil bacteria are effective in mitigating the toxic effects of DMDS on SK-N-MC cells.

4. Conclusion

This study aimed to develop the eco-friendly skid mark scratcher (SMS) combined with bacteria that removes the toxicity of skid marks. For this end, the SMS was composed of the waste materials that are widely discarded in Singapore and are causing environmental problems, such as clam shells, coconut fibers and cigarette butt filters. The skid mark simulated on the cement-based road was removed by cigarette butt filters- or coconut fibers-based SMS, and that skid mark fragments were attached on the SMS remnants, such as sawdust. These SMS remnants, along with the attached skid mark fragments, were then easily collected using a vacuum cleaner, which efficiently removed the residues. This process ensured that the remaining skid mark fragments and SMS remnants were fully cleared from the road surface without leaving visible traces of the skid mark. SMS developed in this study was durable and effective in removing skid marks at least from cement-based road. These findings support the hypothesis that natural and recyclable materials can be used for environmentally friendly skid mark removal.

Skid marks contain various chemical substances, including those like ZnO and DMDS, which are known to exhibit biological toxicity. Therefore, even when skid marks are physically removed, there is a possibility that harmful chemicals may remain on the road surface. This study additionally investigated whether the toxicity of residual harmful chemicals left after skid mark removal could be mitigated by using soil bacteria or *Rhodococcus* sp., which are known for their ability to reduce the toxicity of certain pollutants. In the cell viability experiments, *Rhodococcus* sp. significantly increased the viability of A549 cells treated with 0.15 mM ZnO or 0.032 % DMDS, or the viability of SK-N-MC cells treated with 0.032 % or 0.016 % DMDS. Therefore, this study confirmed the potential of soil bacteria, including *Rhodococcus* sp., in reducing the toxicity of substances like ZnO and DMDS. Soil bacteria showed limited effectiveness in reducing the toxicity of ZnO or DMDS to A549 cells and were similarly ineffective in mitigating ZnO toxicity for SK-N-MC cells. However, both *Rhodococcus* sp. and soil bacteria demonstrated effectiveness in reducing DMDS toxicity for SK-N-MC cells, suggesting their potential utility in specific contexts for DMDS degradation. This study concluded that natural and recyclable materials, such as clam shells, sawdust, sand, coconut fibers, or cigarette filters, are effective for eco-friendly removing skid marks in F1 racetracks.

Further research should investigate the underlying mechanisms of bacterial biodegradation of ZnO and DMDS. Additionally, exploring other bacterial strains or consortia that may have enhanced capabilities in degrading these compounds could provide more comprehensive solutions.

Works Cited

1. Formula 1., Formula-1-2023-Impact-Report
<https://corp.formula1.com/wpcontent/uploads/2024/04/Formula-1-2023-Impact-Report.pdf> (2023)
2. Heideman, G., Datta, R. N., Noordermeer, J. W. M., & Van Baarle, B. (2005). Influence of zinc oxide during different stages of sulfur vulcanization. Elucidated by model compound studies. *Journal of Applied Polymer Science*, 95(6), 1388–1404. <https://doi.org/10.1002/app.21364>
3. ICSC 1586 - DIMETHYL DISULFIDE. (n.d.).
https://chemicalsafety.ilo.org/dyn/icsc/showcard.display?p_card_id=1586&p_version=1&p_lang=en
4. Ramasamy, B., Wu, H., & Yeung, M. (2021). Hosting annual international sporting events and tourism: Formula 1, golf or tennis? *Tourism Economics*, 28(8), 2082–2098.
<https://doi.org/10.1177/13548166211029053>
5. Miller, T. (2016). Greenwashed sports and environmental activism: Formula 1 and FIFA†. *Environmental Communication*, 10(6), 719–733. <https://doi.org/10.1080/17524032.2015.1127850>
6. Ranadev, P., Revanna, A., Bagyaraj, D. J., & Shinde, A. H. (2023). Sulfur oxidizing bacteria in agro ecosystem and its role in plant productivity—a review. *Journal of Applied Microbiology*, 134(8). <https://doi.org/10.1093/jambio/lxad161>
7. Mohammed, S. A., Aman, S., & Singh, B. (2024). Unveiling the positive impacts of the genus *Rhodococcus* on plant and environmental health. *Journal of Experimental Biology and Agricultural Sciences*, 12(4), 557–572. [https://doi.org/10.18006/2024.12\(4\).557.572](https://doi.org/10.18006/2024.12(4).557.572)
8. Kitagawa, W., & Hata, M. (2023). Development of Efficient Genome-Reduction Tool Based on Cre/loxP System in *Rhodococcus erythropolis*. *Microorganisms*, 11(2), 268.
<https://doi.org/10.3390/microorganisms11020268>
9. Chang, N. (2021, January 6). No ifs or butts: Cigarette rubbish is littering Singapore’s coastline. *CNA*.
<https://www.channelnewsasia.com/singapore/single-use-plastic-marine-trash-cigarette-butt-1323081>
10. Bondarev, A., & Gheorghe, C. (2022). The use of lignin for environmental protection: An overview of recent literature. *The Annals of “Dunarea De Jos” University of Galati Fascicle IX Metallurgy and Materials Science*, 45(1), 5–8.
<https://doi.org/10.35219/mms.2022.1.01>
11. Liu, Z., Lv, X., Xu, L., Liu, X., Zhu, X., Song, E., & Song, Y. (2020). Zinc oxide nanoparticles effectively regulate autophagic cell death by activating autophagosome formation and interfering with their maturation. *Particle and Fibre Toxicology*, 17(1).
<https://doi.org/10.1186/s12989-020-00379-7>
12. Api, A., Belsito, D., Biserta, S., Botelho, D., Bruze, M., Burton, G., Buschmann, J., Cancellieri, M., Dagli, M., Date, M., Dekant, W., Deodhar, C., Fryer, A., Gadhia, S., Jones, L., Joshi, K., Kumar, M., Lapczynski, A., Lavelle, M., . . . Tokura, Y. (2021). RIFM fragrance ingredient safety assessment, dimethyl disulfide, CAS Registry Number 624-92-0. *Food and Chemical Toxicology*, 149, 112122.
<https://doi.org/10.1016/j.fct.2021.112122>

Independent Research Project on Affect Theory in Kathak Performance

By Aditi Dhupkar

Abstract

The classical Indian dance form of *Kathak* has migrated from its normative live performance settings to its debut on the Bollywood screen. These two methods of presenting a Kathak dance contain highly different material and generate new reactions from the audience. In this paper, I identify and examine the differences between pre-recorded and cinematic representations of *Kathak* and live *Kathak* performances. I analyze the ways that performance context alters the relationship between dancers, musicians, and their audience.

This paper uses three Bollywood scenes and their song and dance performances as case studies to explore their integration of traditional dance and musical aesthetics. Applying the concepts of *rasa*, *bhava*, *abhinaya*, and *jugalbandi*, I investigate how *Kathak* dancers develop affective environments and create a relationship of intimacy with their audiences (Nimjee). Using an interview with *Kathak guru* Dr. Kriti Rakesh and drawing from my own *Kathak* dance experience and training, I explore how interactions between musicians and dancers in the live performance setting can circulate affect and emotions with their audience. This study also investigates the stigma of *Kathak* in Bollywood cinema. My research is contextualized through a review of literature within ethnomusicology, dance studies, and film studies, providing a deeper understanding of how effective methods are employed to garner audience intimacy. I raise the following questions in this study: Does choreographing emotive gestures such as facial expressions and bodily movements to the rhythm of the downbeat help communicate and develop a relationship with audiences? How has the evolution of *Kathak* affected the social status and the perceptions of *Kathak* dancers? I will answer these questions through an exploration of existing literature, through three case studies from Bollywood cinema, including the films *Devdas* and *Bajirao Mastani*, the television series *Heeramandi*, and through an ethnographic interview with an experienced Kathak dancer and dance guru (teacher), Dr. Kriti Rakesh.

My research is significant to the study of *Kathak* dance, providing an overview of current literature, providing context for historical performance practice, and analyzing the implications of moving a traditionally live art form to a glamourized cinematic context in Bollywood. This study considers the stigmas against the social status of *Kathak* dancers and the value of the form, namely that it proves those stigmas to be unwarranted and contextualizes affect theory in the performing arts.

Paper

My own experience as a Kathak dancer has led me to many questions that compare Bollywood cinema's depiction of *Kathak* in contrast with contemporary live *Kathak* performances. *Kathak dance guru* (teacher), Dr. Kriti Rakesh, provided valuable expertise and insight related to managing a live performance, and maintaining relationships between dancers,

musicians and audience members. Also, he helped answer outstanding questions I had related to the historical practice of the dance form. Rakesh explained, "Bollywood is a big melting pot of whatever dance form there is on Earth. Amongst those many forms, one stands out; *Kathak*." *Kathak* is one of the eight classical dance forms from India. The earliest text with roots in *Kathak* is the *Natya Shastra*, a book detailing the performing arts written between the 2nd century BC and the 3rd century AD. *Kathak* was believed to be founded during the *Bhakti* Movement (Massey, 6-7). Not only has *Kathak* influenced the Bollywood dance form through *ragas* for musical structure, *mudras*, and facial movements, but it has been present in Bollywood cinema as well. This paper presents an analysis of the connectivity between the *Kathak* dance form and Bollywood depictions of this form. By comparing live and on-screen performances of *Kathak* dance, I examine the relationships between dancers, musicians, and their audience, and argue that live performances are a distinctly different form than those presented in Bollywood cinema. I support these claims through an investigation of the elements for fostering audience intimacy used in different case study examples from film and television, through a review of existing literature on the practice of *Kathak*, and through an interview with an experienced *Kathak* dancer and master *guru* (teacher), Dr. Kriti Rakesh.

This study analyzes the representation of *Kathak* dance in Bollywood cinema using three scenes and their song and dance performances as case studies, exploring their integration of traditional dance and musical aesthetics. It aims to contextualize these depictions within ethnomusicology, dance studies, and film studies, providing a deeper understanding of how affective methods are employed to garner audience intimacy. This paper will also explore how affect theory is employed in live performances of *Kathak* dance. I apply affect theory as a methodology of analysis to investigate *Kathak* dancers' perspectives on their own practice and relationship with the audience (Nimjee). I raise the following questions in this study: Do emotive gestures such as facial expressions and bodily movements to the rhythm of the downbeat help communicate and develop a relationship with audiences? How has the evolution of *Kathak* affected the social status and the perceptions of *Kathak* dancers? I will answer these questions through an exploration of existing literature, through three case studies from Bollywood cinema, including the films *Devdas* and *Bajirao Mastani* and the television series *Heeramandi*, and through an ethnographic interview with an experienced *Kathak* dancer and dance *guru* (teacher), Dr. Kriti Rakesh.

Kathak dancers and musicians generate affect, or the product of the exchange of powerful emotions, through their performance. Drawing from Ameera Nimjee's work on *Kathak* dance and affect, I suggest these emotions are shared with their audience in both live *Kathak* dance performances and in *Kathak* dance scenes in Bollywood cinema. I investigate their methods of developing and sustaining audience intimacy through camera angles, facial expressions, costuming, and glamour. Through my own experience as a *Kathak* dancer and through an ethnographic interview with my *Kathak* dance *guru* (teacher), Dr. Kriti Rakesh, I explore the importance of the onstage relationship between dancers and musicians through an analysis of rhythmic communication, musical accents, bodily gestures, and competition of skill between

performers. I will compare depictions of *Kathak* in Bollywood to live *Kathak* performances and determine how differences in setting and context replicate affect through performance for an audience. Some important terminology used in this paper includes the term *bhava*, an emotion that a performer attempts to convey. *Rasa* is the reaction evoked from the audience in response to a performer's *bhava*. *Rasa* and *bhava* are closely related to affect, because they serve as the foundation for emotional expression in *Kathak*. Affect theory scholar Sara Ahmed describes how emotions circulate through bodies, and this movement accumulates and generates new affect (Ahmed). I define *lehra* in the same way that the scholar Ameera Nimjee does in her article titled *Playing Dance and Dancing Music: The work of Intimacy in Kathak*, as an "ostinato melody that marks the beats within a rhythmic cycle" (178). Describing and analyzing a *Kathak* performance requires knowledge of the music, or the *lehra*. The term *bol* describes the spoken translations of drum music. Lastly, I define *abhinaya* as the "acting" aspect of *Kathak*. *Abhinaya* is important to show the roots of *Kathak* as a form of storytelling in India years ago, when there was no means for written messages. Instead, *Kathakars*, a professional class of storytellers, used to perform epics and mythological stories with the added element of *abhinaya* (acting) for others (Kothari, 1). *Kathak* has evolved extensively in its modern form, but storytelling is still a fundamental part of performance. These terms will be used to better understand the relationships between *Kathak*, Bollywood, affect theory, and developing audience intimacy.

Kathak dancers all learn from their *guru* (teacher). Although this form of dance was originally used for storytelling, over time, it took many other forms, including that of a devotional dance, a courtesan's dance, and eventually into the form of performing arts that it is today. Though it has gone through many social stages, the stigma of being a courtesan's dance still sticks in *Kathak* today. Oftentimes, this deters people from learning this dance form (Rakesh). However, *Kathak*, in its current form, does not limit participation to a certain role or demographic. It is a form of dance like any other, open to the general public. A *guru* can be anyone who has obtained a diploma in *Kathak*, which is attained by passing seven levels of *Kathak*. I am in the sixth level, *Visharad Pratham*, myself. At Aakriti Kathak Academy, the dance studio I study with, classes usually include the teaching of how to dance and write compositions, information about *taals*, different forms of *abhinaya* including *ghatbaav* (stylized walking) or *navikas* (heroines), and the history of *Kathak* and its important figures.

Indian classical music holds an equal emphasis on singing, dancing, and playing instrumental music. There are four kinds of instruments in Indian performance culture: *Tata Vadya* (string instruments), *Avanadha Vadya* (percussion instruments), *Ghana Vadya* (brass instruments), and *Susir Vadhya* (wind instruments). These are important to dance, because dancers and singers rely on the tempo from these instruments, and these instruments provide music and life for a performance (Dudhe, 246-248). Instrumentalists usually sit on the side of the stage, perpendicular to the dancer(s). Instrumentation in *Kathak* typically includes the *sitar*, *tabla*, *pakhawaj*, and the vocalist (Rakesh). The musicians play different *taals* (background loops) on which compositions take place. *Abhinaya* compositions can also have songs, which require more specific music laid on a *raga* with a vocalist singing the lyrics. Anyone can be an

instrumentalist. However, some technical aspects require more training in music, such as *jugalbandhi*.

Scholar Ameera Nimjee focuses not only on the relationship between musicians and dancers, but also on how this relationship generates emotions and creates bonds with the audience. These bonds are products of emotive generation referred to as affect. Many scholars have researched affect theory, and how it influences relationships even though it is not a tangible object. Affect theory has multiple definitions. Affective studies can come in the form of building bridges between the humanities and intensive sciences or serve as an analysis of psychology without a conclusion (Figlerowicz, 3). Affect theory has to do with emotions, and how these emotions produce connections between people. In music, affect theory is described as "a set of conceptual tools and key terms for thinking through music in its non semantic, pre semantic, or extralinguistic registers" (Robinson et. Nickelson, 2). There are many ways that affect can be shared among those present in a live performance context. Christopher Small's theory of *Musicking*, which describes effective potential and active participant value to both performers and audience members, can be applied to the context of a live *Kathak* performance. According to Small, "All those present, both performers and listeners are engaging in the encounter, and all are contributing to its nature through the relationships they establish with one another during performance" (2). Small's theory posits that everyone, including the listeners, participates in the making of music. Small argues that live performances evoke emotions that create intimate bonds between everyone attending. This is one way that effect is measurable in live performances, while this cannot happen on-screen. The nature of the film screen is such that it limits the interactive abilities of the audience. Audiences and performers begin to form intimacy with one another through sharing emotions or being in the same affective state through the medium of performance. In a special edition article, scholar Lauren Berlant defines intimacy: "It can be portable, unattached to a concrete space: a drive that creates spaces around it through practices. The kinds of connections that impact people, and on which they depend for living (if not "a life"), do not always respect the predictable forms..." (284). Intimacy produced through performance provides a way for performers to harness affect and engage more deeply with their audience during a live performance. Dance scholar Ameera Nimjee analyzes how intimacy plays directly into *Kathak* dance. Nimjee uses the phrase "the work of intimacy" to describe the relationship between the dancer and their musicians in *Kathak* (176). This shared intimacy can be displayed through glances, gestures, and footwork, and helps engage the audience and emphasize details and technicality achieved through the hours of practice of the performers. The work of intimacy can only be achieved by those proficient through years of training in their field; unlike the dance itself, this intimacy with other performers and their audiences cannot be choreographed, and performers must be ready to implement new bodily gestures and interactions on the go (Rakesh).

Intimacy is developed between performers and with audiences in many ways in the practices of *Kathak*. One example of this intimacy is showcased in *Kathak* through *jugalbandi*. *Jugalbandi* is when a musician and a dancer play a pattern, and then repeat it back to the other in

their own language (Nimjee, 179). This allows both performers to interact and create a special relationship, while also showcasing their ability to play or dance difficult patterns. This quick back and forth can create a sense of awe and intrigue amongst the audience, as they see how music and dance mirror each other and watch the interactions of both performers. Another example of this intimacy is *sam*, or the downbeat of a rhythmic cycle. Upon reaching the end of a composition, both the dancer and musician emphasize *Sam*, almost as if to tell the audience, “Look, I made it!” when they reach the downbeat on time (Nimjee, 179). This emphasis, along with emotive gestures such as eyebrow raises, head movements, or hand gestures helps the audience understand not only where a composition ends, but also allows the cohesiveness of the music and the dance to work together to unify the audience and performers in sentiment and effort. Live performances of *Kathak* generate affect and information to help create shared experiences between performers and the audience. For example, a *Kathak* dancer counting a *farmaishi chakradhar paran* (specific patterned composition born from the *bow* of the *pakhawaj*, a traditional Indian drum) could use head nods to emphasize when the pattern correctly falls on *sam*. This would show the audience how a *farmaishi* pattern works and create a point of intrigue for the audience. In this way, a live performance can use actions such as emphasis or bodily gestures to engage and aid the audience using affect in ways that cinema cannot.

In the context of *Kathak*, affect might be best understood through the terms *rasa* and *bhava*. As mentioned above, *bhava* is the emotion that a performer is attempting to convey, while *rasa* is the reaction evoked from the audience in the response to said performer’s *bhava*. The required *rasa* in *abhinaya* (the storytelling or “acting” aspect of *Kathak*), in turn, evokes *bhava* from the audience and generates a new ambience, or the atmosphere where the performance takes place (Rakesh). This ambience changes the audience’s perspective of the performance, even though it isn’t a tangible substance. Like Small’s theory of *Musicking*, *Kathak* performers can facilitate and create a context where the audience plays an essential and participatory role in the dance form. For example, in a concert of Hindustani music, it is normal for the audience to contribute appreciation through clapping or verbal affirmation for the performer (van der Meer, 170). This is also typical in a live *Kathak* performance setting. The added level of audience and performer interaction helps the audience take part in producing an effective atmosphere around the performance. I argue that the ambience of a live performance, including elements such as bodily gestures, live music, and implied emphasis, constitutes a definitively separate experience from that of a recorded performance, such as those I will examine in case studies below.

One of Bollywood’s most prominent features is its song and dance. Song and dance in Bollywood films are hooks, or earworms, a small way for a watcher to recall the movie after watching it. They also became ways to help form connections between people. According to Sangita Gopal and Sujata Moorti in their book *Global Bollywood*, “This trend of Bollywood films being driven by their songs began with the film *Awaara* and its title track” (16-17). Thus, many scholars tend to focus on the song and dance aspect of Bollywood in their papers. Moreover, these scholars emphasize the diversity found in Bollywood dance (Gopal and Moorti, 1). Bollywood has been heavily influenced by multiple styles of dance across the world, ranging

from *Kathak* to the Trinidadian form of *chutney* to the Greek form of *indoprepi* (Gopal and Moorti, 1).

Recent Bollywood films have often included scenes of *Kathak* dance. In the movie *Devdas*, the actress Madhuri Dixit dances *Kathak*-inspired choreography in the song, *Maar Dala*. The choreography isn't the same as the pure *Kathak* seen in a live performance but draws heavy inspiration and uses similar movements. In this scene (scene A), Madhuri Dixit, or Chandramukhi, is depicted as a courtesan. The *Kathak* dance form emerged from what is referred to as *nautch*. *Nautch* was performed in the courts of North and East India from the 1600s to the 1870s during British rule. These performers were women known as *tawaifs* (Chakravorty, 116). The depiction of *Kathak* dancers as courtesans in Bollywood cinema is very frequent. These scenes vary tremendously from live performances of *Kathak*; they lack the human aspect of interaction and in-person effect of emotive gestures, but instead include excessive opulence and camera closeups to make the scenes attractive.

This first scene is covered in sparkling diamonds, drawing in the attention of the audience through spectacle and glamor. Chandramukhi is in love with Devdas, played by Shah Rukh Khan. In this scene, Chandramukhi is dancing to woo Devdas. The camera follows Chandramukhi closely, focusing on her facial expressions. This creates a feeling of intimacy where the audience sympathizes with her perspective, producing a different form of intimacy than what is generated during a live performance. Chandramukhi, being a courtesan (mistress), “sings” for herself in this scene. This is different from a live performance, where a separate musician would perform this role. She wears a green *anarkali* (dress) with lots of embroidery and a *dupatta* (scarf) covering her head. This adds to the glamorous aspect of the scene and creates an aspirational affinity within the audience, where the viewer admires the beauty and wants more. This differs from the intimacy that performers work to create in a live performance of *Kathak* using bodily gestures and communication among performers.

The music in this scene includes the noise of the bells, or *ghungroos*, that *Kathak* dancers tie around their ankles, and the traditional *tabla* (drum) also plays in the background. However, the musicians are out of frame and cannot be seen. This takes away from the aspect of intimacy that is usually performed during a live performance of *Kathak* between the dancer and their musicians (Nimjee). The camera's primary emphasis remains on the dancers, although Chandramukhi is the central focus. The background dancers wear different colored costumes and do more complex choreography rather than just facial expressions. This is unusual, as this only happens in live performances if the dancers are attempting to portray a specific character. This helps incorporate more *Kathak*, while at the same not making the dance too difficult for the actress starring in the film, who is on most occasions not trained in any classical dance let alone *Kathak*. The camera also follows the main character, Devdas, or Shah Rukh Khan. Since this scene depicts their first meeting, the camera closely follows both character's facial expressions and interchanged glances. Some of the steps also follow the meaning of the song lyrics, emphasizing the *abhinaya* (acting) aspect of *Kathak*. Overall, in this scene, they attempt to produce intimacy through camera work and emotive gestures such as facial expressions or bodily

movements. Since a film cannot create the same atmosphere as a live performance, filmmakers use extravagant clothing and sets to convey ideas. This helps not only maintain audience attention, but also creates an image of this place as “paradise on Earth” that is a driving point for many Bollywood films and TV series to help engage the audience and make them feel emotions of amusement, diversion, and entertainment (Inden, 495).

The second scene (scene B) from the film *Bajirao Mastaani* also takes place in a royal court. In this song called *Mohe Rang Do Laal*, excessive opulence in a grandiose and ornate palace setting, and strategic use of color that creates contrast between the main character and backup dancers are showcased in this scene. The main character, Peshwa Bajirao, portrayed by Ranveer Singh, is a *Maratha* king. The other main character, Mastaani, portrayed by Deepika Padukone, is a warrior princess, but she is also a courtesan in Bajirao’s court. In this scene, the grand opulence of the palace is on display. The camera begins with an extremely long shot of the palace before panning in the fountains, and showing the king, Bajirao, in all his glory. This type of panning camera work is used to show large size or majesty (Heiderich, 7). Glamor in the scene serves as a “wow” factor and helps get the attention of the audience. After this, the tinkling of *ghungroos* (bells) begins once again, signaling the entrance of the dancers. The main character is in a brighter color than the backup dancers, to highlight her more. Dressed in bright white, she wears a heavy *anarkali* (dress) with her *dupatta* (scarf) covering her head. This outfit replicates the traditional dress, where women typically covered their heads. While the backup dancers have slightly more complex steps, Mastaani sits down, and the camera zooms in for a close-up shot of her facial expressions, which is used to create viewer intimacy and sympathy towards the character’s actions (Heiderich, 7). This allows the audience to feel the character’s emotions and generates intimacy.

In addition to this, her hands are colored with red pigment called *alta*, to highlight any hand movements she does and frame her face. The dancers also wear *anarkalis* (dresses) that flare out to showcase their *chakras* (spins). These smaller details enhance the performance, create a sense of wanting in the audience, and make it more fun to watch. Female musicians are pictured playing the *sitar* in this scene. Unexpectedly, the main character “sings” for herself. This is unusual, as a performer would have a separate musician to sing for them in a live *Kathak* performance. However, courtesans in the past used to sing for themselves (Chakravorty, 116). As expected, the main character’s jewelry is covered in diamonds. Mastaani also uses many *mudras* (hand gestures) that are from *Kathak*. The camera follows the developing relationship between the two main characters, making sure to capture their interchanged glances and gestures. This helps generate a sympathetic effect with the audience, even if the interaction is not directed towards them.

Some of the steps also follow the lyrical meaning of the song. The song is speaking of the color red, and how this color plays into blushing. This meaning is reflected in the *laal* (red) color on Mastaani’s hands. It also talks about the relationship between the two main characters. There is a *sargam* (when the musical notes are sung instead of played) break in the song, and all of the dancers cluster together to do more quick-paced, complex steps. It is typical for slower *Kathak*

compositions to have small, faster footwork sections, and this is reflected through the *sargam* section in this song. The music in this part is sung by unseen musicians, and the lyrics are drawn from the *bol* that the ghungroos are producing.

The third case study (scene C) from the TV show *Heeramandi* is a song called *Phool Gendwa Na Maaro*. This scene showcases a courtesan named Bibbojaan, who is portrayed by Aditi Rao. She is performing her last dance, which is why there are many *nawabs* (royalty) around. Once more, a *Kathak* dancer on-screen is regarded with a low social ranking. Her *nawab* is showcased right at the front, throwing money towards her. The scene starts by touring the audience through the grand house. There are many water fountains, sparkling chandeliers, and elaborate designs in the architecture. This creates a sense of awe for the viewer, also known as *adhbhuta rasa* (the *rasa* of awe / wonder). Bibbojaan is dressed in dark red to mark the sad occasion of her final performance, but the *anarkali* (dress) is still heavily embroidered, she is laden with jewelry, and her *dupatta* (scarf) covers her head for modesty. In the background, the other dancers wear darker brown, to highlight Bibbojaan's role as the main character. They do quick spins and *Kathak* choreography, while Bibbojaan uses smaller *mudras* and gestures to interact with her *nawab* and generate affect within the audience.

Though there is *sitar* and *tabla* playing in the music, they are not picturized on the screen. Like the previously mentioned case studies, Bibbojaan is pictured to be “singing” for herself in the scene. Throughout the scene, Bibbojaan and her *nawab* have an established intimacy by exchanging glances and through Bibbojaan's bodily gestures, while in exchange, the *nawab* showers her with money. In this scene, there is no sung *sargam* section; instead, there is a *jugalbandi* between the sitar and the tabla. During that section, the choreography also becomes more complex, which functions to “wow” the audience and retain their attention. After this, the music slows back down and the camera focuses more on both the *nawab* and Bibbojaan's expressions, showing the sadness and intimacy on their faces.

All of these scenes are independently different, but they hold a lot of similarities. In all three of these scenes, the *Kathak* dancer is portrayed as a courtesan. Though *Kathak* was a courtesan's dance, this was years ago during British rule and the Gupta Regime (Dudhe, 184). *Kathak* has also been used for religious reasons. This depiction creates a stigma towards *Kathak* in real life, and even drives families away from wanting their kids to learn *Kathak* (Rakesh). The Bollywood depiction of *Kathak* is also highly different; firstly, it comes from a different historical perspective. In Bollywood-*Kathak* portrayals, the main dancers are often singing for themselves, which is unusual for a live performance. Costumes are also more heavily embroidered on screen, and they are more traditionally conservative. For example, they are all long-sleeved, and the *dupattas* (scarves) are pinned to cover the performer's heads. Bollywood also uses different tactics to generate affect with the audience and maintain their attention. In scene A, there was an emphasized focus on the main character's emotions. In a live performance, full-body gestures and *mudras* (hand gestures) would be used to display emotions along with facial expressions. However, in scene A, the camera follows the main character's expressions, interacting with Chandramukhi. This changes the way that the audience sees her (Heiderich, 6).

Thanks to the close-up camera angles, the audience can feel Chandramukhi's emotions more clearly. Moving from long to close-up camera shots, like in this scene, helps depict intimate emotions (Heiderich, 6). These emotions, even if not directed at the audience, generate affect through creating an intimate connection between characters. Another way that these movies attempt to recreate the atmosphere of a live performance is through creating *adbhuta rasa* (feeling of awe) within the audience. This is showcased in scene B. This scene begins by touring the opulent palace and captures the audience's attention. The main characters' elaborate clothing and jewelry contribute to this grandness, evoking affect in the audience through *adbhuta rasa* and a feeling of desire. The last way that Bollywood cinema tries to recreate audience intimacy generated at a live performance is through character relationships. Characters build relationships with each other, since they cannot interact with the audience through the film screen. This generates an atmosphere of affect within the audience. In scene C, the main character, Bibbojaan, has many interactions with her *nawab*, who is sitting in front. The two exchange bodily gestures and glances that establish an emotive relationship for the audience to watch. This attracts the audience's attention by creating a sense of intrigue and adoration for the characters and their relationship.

To further explore this topic, I conducted an interview with my *guru* (teacher), Dr. Kriti Rakesh. I have been dancing at her academy for around nine years. She graduated from Creighton University with a PhD in Biomedical science, and now owns a non-profit organization called Aakriti Kathak Academy. Classes consist of teaching how to dance compositions, count compositions, learn the history of *Kathak*, and information about *taals*. Rakesh also plans recitals once or twice a year to give her students practice performing in front of a crowd. She has received her diploma in *Kathak* twice; once in her youth, and once after she realized she wanted to pursue *Kathak*. The process of data collection for this interview included recording and transcribing our conversation. This interview was conducted face-to-face, carried out using a romantic approach. Within qualitative interviewing methodology, Welch and Piekkari define a romantic approach to an interview as an argument "that interviewers need to engage with interviewees on a personal level, in order to establish an atmosphere of rapport and trust that will allow interviewees to produce genuine and open responses" (420). Engaging in discussion over the work of intimacy in live performances, I asked Rakesh questions about her experience with her musicians and interactions with the audience. I investigated her feelings towards the importance of affect in *Kathak*. She believes that affect theory is mostly conveyed through *abhinaya* in *Kathak*. However, she also believes there are many other aspects of a live performance that produce an intense atmosphere between all the people present in the experience.

When describing the affective nature of musical vibrations in a live performance, Rakesh describes, "The minute one tabla person and one sitar person and a vocalist will sit, the vibrations that it creates, it's- it's pure joy." And when this joy is created, "that is what the audience appreciates; that, 'Wow, they are just doing this in front of you. No retakes.'" The live interactions of these musicians and their vibrations have the power to create an atmosphere on

their own; but add in the dancer, and this generation of affect becomes even stronger. The *Kathak* dance form is highly interactive. According to Rakesh, “more than other dance forms, *Kathak* is more interactive and informal, so the style of presentation is- you know, the dancer is always interacting with their audience as well as the musicians, and it's not very set. And choreographed.” This is further reflected in how *Kathak* dancers and musicians often make up compositions or patterns. Rakesh describes, “The dancer may come up with something spontaneous, and throw it towards the musician, and the musician then has to respond, or other way around, the musician may come up with something, and the dancer can respond in their own way, and, you know, sometimes they- they involve the audience in this process as well, and that way it becomes a circle. Of the dancers and the musicians and the audience. And that way it becomes very interactive.” When the audience is watching this, they know that it is all in one take.

Rakesh highlights and describes the differences between live and cinematic performances of *Kathak* dance. She believes that even if the audience isn't consciously aware, they know that a movie has had multiple takes, retakes, and camera angles from all sides. Thus, it is more difficult for the hard work to shine through in Bollywood cinema. On stage, however, “They [the audience] can appreciate it; they don't have to be experts in that field, but they can appreciate the- the level of training and expertise the team has put in to be at that point.” However, on the downside, there is only one chance for this magical sensation to be created. And if someone hasn't correctly done their work, then a performance can fall apart. “So even if one person has not done their 'homework', they'll pull down the entire hard work of every single person on that team.” Live performance creates a higher pressure for performers, but also generates greater emotion and affect from the audience. My *guru* also strongly believed that *Kathak* in Bollywood cinema is not accurately represented. According to her, “*Kathak* in cinema is stuck in the past”. *Kathak* dancers are not depicted as respectful artists, but instead as lowly courtesans. There are few movies where a *Kathak* dancer has gained a “respectful lead role.” This has caused a stigma against families wanting their children to learn *Kathak* even today.

Analyzing the representation of *Kathak* dance in Bollywood cinema using three scenes and their song and dance performances as case studies, this paper has explored their integration of traditional dance and musical aesthetics. I have used my own experience as a *Kathak* dancer, along with an interview of my *guru* (teacher), Dr. Kriti Rakesh, to examine the relationship between onstage performers and the audience. I have answered ancillary questions, including how the evolution of *Kathak* has affected social stigmas against the dance form today and how emotive gestures help create communication between the audience and performers. Creating onstage relationships during live performances of *Kathak* using emotive gestures generates an affective atmosphere. This atmosphere helps engage the audience in a performance. A relationship can be created between performers using emotive gestures such as bodily movements, *mudras* (hand movements), or through interchanged glances and facial expressions. Live performances also have the vibrations of music and lyrical meaning that aid the production of an affectively charged atmosphere. *Kathak* has also been presented through the film screen.

However, this method of generating effect isn't applicable to this setting of presentation. In Bollywood cinema, filmmakers use a variety of techniques, including camera shots, character relationships, emphasis on smaller gestures and expressions, and excessive opulence. These strategies are used as an attempt to generate effect through the screen and capture the audience's attention. Past stigmas related to *Kathak* and its historical place in society are also visible in film depictions of dance scenes. *Kathak* performers are often presented as courtesans of low status in Bollywood cinema. This has created an impression in people's minds that contributes to families not wanting their kids to participate in this dance form. While the dance form is not relegated to a specific role or social status today, Bollywood representations may contribute to lingering residues from the past related to the role of courtesans in *Kathak*. However, in live performance settings today, *Kathak* dancers are respected artists.

This research paper has addressed the following questions: How do *Kathak* dancers develop affective environments and create a relationship of intimacy with their audiences? Do emotive gestures such as facial expressions and bodily movements to the rhythm of the downbeat help communicate and develop a relationship with audiences? How has the evolution of *Kathak* affected the social status and the perceptions of *Kathak* dancers? *Kathak* dancers use bodily gestures and facial expressions to develop effective environments. These environments, in turn, create emotively charged bonds with the audience. Some of these bodily movements and gestures, such as head nods, or facial expressions, such as eyebrow movements, can communicate information to the audience, including when a composition ends or signal the pattern of a composition. The evolution of *Kathak* has created a stigma against the social status of this form in today's society, but *Kathak* and its performers still remain highly regarded in the performing arts world. I have answered these questions through an examination of three case studies in Bollywood films, an ethnographic interview with *guru* Dr. Kriti Rakesh, and personal experience as a *Kathak* dancer. Though the focus of this paper only looked at three case studies, dance and music scholars find significant differences between current live *Kathak* performances and those depicted in Bollywood. By examining the affective tools used by both modes of *Kathak* performance and identifying their differences in audience response, we can conclude that they are different art forms, though they derive from the same historical source. Opportunities for further research on Indian dance forms such as *Oppari* provide another outlet to explore the affective relationship between performers and audience members. *Oppari* is a form of weeping songs usually performed by women in grieving or professional males of the *Harijan* caste community. These songs can also be performed at home and are used to showcase feelings of lament or display problems. Thus, *Oppari* helps give women a voice to speak out when they are mostly limited in their actions based on their caste. *Oppari* can be used to speak out against injustices in society (Greene). The emotional ties that come with *Oppari* are very powerful and generate an effect with a sympathetic or driven audience for change. When performed correctly, *Oppari*'s strong emotions of sadness can help generate effective charges that have impacts on social issues or the performer's problems. This effect is produced through emotive expression and meaningful song lyrics. Through many dance styles, affect can be used to help gain the

audience's attention, allow them to understand and enjoy a performance, or even emotionally empower them to take action. In *Kathak*, this effect is generated through onstage relationships made of emotive expressions, bodily gestures, and interchanged glances between dancer and musician. However, there are many ways that effect is produced on the Bollywood screen in a *Kathak* performance, including excessive opulence, character relationships, and camera shots. The study of affect theory in the performing arts can create more enjoyable and simple performances for an audience.

Works Cited

- Aishani M. Nanda. "The Commercialisation and Commodification of Bollywood Music." *International Journal For Multidisciplinary Research*, vol. 6, no. 2, Mar. 2024, p. 14562. DOI.org (Crossref), <https://doi.org/10.36948/ijfmr.2024.v06i02.14562>.
- Ahmed, Sara. "Affective Economies." *Social Text*, vol. 22, no. 2 (79), June 2004, pp. 117–39. Silverchair.
- Arora, Kamya. *Kathak: As Employed by Sanjay Leela Bhansali in Three Key Works*. 2017. *Audience Response and Expressive Pitch Inflections in a Live Recording of Legendary Singer Kesar Bai Kerkar | Expressiveness in Music Performance: Empirical Approaches across Styles and Cultures | Oxford Academic*.
<https://academic.oup.com/book/27746/chapter-abstract/197948491?redirectedFrom=fulltext>. Accessed 7 Sept. 2024.
- Berlant, Lauren. "Intimacy: A Special Issue." *Critical Inquiry*, vol. 24, no. 2, Jan. 1998, pp. 281–88. DOI.org (Crossref).
- Chakravorty, Pallabi. "Dancing into Modernity: Multiple Narratives of India's Kathak Dance." *Dance Research Journal*, vol. 38, no. 1–2, July 2006, pp. 115–36. Cambridge University Press.
- Chakravorty, Pallabi, and Nilanjana Gupta. *Dance Matters: Performing India*. Routledge, 2010.
- Dudhe, Vaishali. *Dance Trail... Exploring the Heart of Kathak*. Manjiri Nrityakala Mandir. 2015.
- Figlerowicz, Marta. "Affect Theory Dossier: An Introduction." *Qui Parle*, vol. 20, no. 2, 2012, pp. 3–18. JSTOR.
- Greene, Paul D. "Professional Weeping: Music, Affect, and Hierarchy in a South Indian Folk Performance Art." *Ethnomusicology OnLine*, no. 5, Jan. 1999. [go.gale.com](https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=10927336&v=2.1&it=r&id=GALE%7CA229301667&sid=googleScholar&linkaccess=abs),
<https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=10927336&v=2.1&it=r&id=GALE%7CA229301667&sid=googleScholar&linkaccess=abs>.
- Heiderich, Timothy. *Cinematography Techniques: The Different Types of Shots in Film*.
- Inden, Ronald. "What Happens in a Hindi Film?" *Asian Journal of Social Science*, vol. 41, no. 5, 2013, pp. 492–509.
- Kothari, Sunil. *Kathak, Indian Classical Dance Art*. Abhinav Publications, 1989.
- Massey, Reginald. *India's Kathak Dance, Past Present, Future*. Abhinav Publications, 1999.
- Nimjee, Ameera. *Playing Dance and Dancing Music | 12 | The Work of Intimacy in Kathak*.
<https://www.taylorfrancis.com/chapters/edit/10.4324/9781351116183-12/playing-dance-dancing-music-ameera-nimjee>. Accessed 20 July 2024.
- Sukhatankar, Ojasi. "Practice and Pedagogy of Indian Classical Dance (ICD): Analysis in Comparison with India's Folk and Bollywood Dances." *The Chitrolekha Journal on Art and Design*, 14 July 2016,
<https://chitrolekha.com/indian-classical-dance-folk-bollywood-dances/>.
- Rakesh, Kriti. Qualitative Interview. 9 Aug. 2024.

- Robinson, Dylan, and Patrick Nickleson. "12 The Feeling of Knowing Music." *12 The Feeling of Knowing Music*, Duke University Press, 2023, pp. 273–92. www.degruyter.com.
- Sahu, Deepshikha. *Kathak Attire Evolution: From Ancient to Modern Era*. 2024.
- Sanyal, Shankha, et al. *Ragas in Bollywood Music - A Microscopic View through Multifractal Cross-Correlation Method*.
- Sen, Biswarup. "The Sounds of Modernity." *Global Bollywood*, edited by Sangita Gopal and Sujata Moorti, NED-New edition, University of Minnesota Press, 2008, pp. 85–104, <http://www.jstor.org/stable/10.5749/j.ctts8hz.5>. JSTOR.
- Small, Christopher. "Musicking." *Wesleyan University Press*, <https://www.weslpress.org/9780819572240/musicking/>. Accessed 7 Sept. 2024.
- "The Captivating Fusion: How Kathak Shapes Bollywood Dance." *Dr. D. Y. Patil School of Liberal Arts, Pimpri, Pune*, <https://liberalarts.dpu.edu.in/blogs/captivating-fusion-kathak-shapes-bollywood-dance>. Accessed 23 July 2024.

The Hidden Crisis: Innovative Strategies for Climate-Induced Distress in Low Socioeconomic Populations

By Riya Kansagra

Abstract

Climate change has a disproportionate impact on the mental health of low socioeconomic (SES) communities, as barriers such as rising temperatures and natural disasters significantly increase anxiety, depression, and PTSD, creating a need for interventions. Therefore, this paper will examine the mental health effects of climate change disproportionately impacting low SES communities in the U.S., and what interventions are being developed to address these disparities. Previous research has shown that low-income populations are more vulnerable to environmental hazards while also having limited access to mental health services, further heightening these issues. Examining the challenges to mental health care in low SES communities reveals key obstacles and highlights the need for both immediate interventions and long-term support strategies. Implementing mobile mental health units, expanding community-based programs and nonprofits, and expanding trauma-informed care in hospitals is key to sufficient recovery. Developing specific mental health initiatives can ensure consistent and ongoing care for low-income individuals in light of climate change, and could maximize the recovery and resiliency of these communities.

Introduction

Climate change is a significant worldwide challenge, affecting ecosystems, health systems, and communities (NOAA). Specifically, rising temperatures and natural disasters threaten economies tremendously (NOAA). As global temperatures rise, the severity of natural disasters increases, leading to further damage. For instance, between 1980 and 2023, the United States experienced over 360 weather and climate disasters, each exceeding \$1 billion in damages, which totaled \$2.6 trillion in economic losses (NOAA). Additionally, rising sea levels pose a significant threat to coastal communities, forcing individuals to be displaced, and creating much economic instability. Moreover, climate change disrupts food production by worsening droughts and floods, which in turn can reduce crops, increase prices, and heighten world hunger (OAR US EPA). Among its many repercussions, climate change has a direct impact on mental health, contributing to stress, anxiety, depression, and other health conditions (Ramadan and Ataallah). The trauma of losing homes, jobs, or even loved ones due to extreme natural disasters is devastating and has been proven to increase the rates of PTSD and suicide effects in populations (Padhy et al.). Furthermore, climate change has caused many social disruptions such as forced migration of families and economic instability, particularly in thus far vulnerable communities (Ramadan and Ataallah). In areas that experience climate-related disasters constantly, individuals can develop stress disorders like anxiety due to the continuous threat to their safety, homes, and families (Ramadan and Ataallah). According to the National Institutes of Health, individuals who survive natural disasters are at a 40% higher risk of developing PTSD compared to individuals who do not (Padhy et al.). Along with PTSD, research has shown that

extreme heat exposure correlates with a 7.9% increase in suicide rates and a 21% rise in mental health-related emergency room visits (Ramadan and Ataallah). Without intervention, the mental health crisis increasing because of climate change is expected to worsen, placing additional stress and strain on healthcare systems.

Along with rapidly increasing climate-related health concerns, they are explicitly growing in marginalized communities, such as racial and ethnic minorities, immigrants, and individuals with disabilities as well as those of lower socioeconomic status (SES) backgrounds, which refers to those with limited financial resources and education. These communities often face many challenges because of higher risks of health problems and limited access to resources (USDA). These overlapping factors create a system of inequality, where climate-related health risks are magnified because of these roadblocks, making it even more difficult for these vulnerable populations of people to access the necessary resources they need.

Low socioeconomic status (SES) communities in the U.S. endure the uncalled-for burden of climate change-related mental health effects due to a combination of environmental, economic, and social factors (Giudice et al.). Environmentally, these communities are often located in areas most vulnerable to climate change, such as floodplains, urban heat islands, and regions with faulty infrastructure (Giudice et al.). The U.S. Department of Agriculture (USDA) records that economically disadvantaged neighborhoods frequently face higher exposure to environmental hazards like air pollution or extreme heat due to the lack of green spaces or cooling systems (USDA). Low SES individuals are more likely to live in areas with higher pollution and extreme heat, which can all circle back to depression and anxiety disorders (Smirnova et al.). Research shows that individuals living in low-income neighborhoods often experience greater exposure to air pollution, such as particulate matter (PM) and ground-level ozone, which are connected to many cardiovascular and respiratory problems (Smirnova et al.). The stress from dealing with the physical effects of pollution, specifically asthma, respiratory infections, and other chronic diseases can heighten anxiety and depression. The Environmental Protection Agency (EPA) has also identified that poor air quality can directly affect the brain, which can trigger neuroinflammation, which may lead to cognitive diminishment and again increased vulnerability to mental health disorders (OAR US EPA).

Moreover, people in low-income communities often struggle with financial instabilities and lack of access to healthcare, which can make it even harder for them to get the mental health support they yearn for and need (Giudice et al.). Additionally, some natural disasters can cause these communities to lose their employment, and have food shortages, further heightening their vulnerability to mental health (USDA). A real-life example of this is the aftermath of Hurricane Katrina, which wrecked New Orleans in 2005. Many low-income residents, particularly Black and Latino communities, were relocated, losing not only their homes but also their jobs (Bienvenu). These disturbances, combined with access to support and resources, left many of these individuals in extreme states of distress and despair (Princeton University et al.). Additionally, the loss of employment and lack of resources, only left the overall state of mental health to worsen, demonstrating how the economic challenges directly affected these

communities. Therefore, this paper will focus on the mental health effects of climate change disproportionately impacting low socioeconomic communities in the U.S., and what interventions are being developed to address these disparities.

Disproportionate burden on low socioeconomic communities

Low socioeconomic communities face a variety of challenges intensified by climate change. These communities often face poverty and poor living conditions (USDA). Poverty is a fundamental issue in these communities as individuals with lower incomes often struggle to afford necessities such as food, stable housing, and medical care (USDA). When extreme weather events occur, these individuals frequently lack the financial resources to rebuild homes, relocate, and recover from lost money, trapping them in economic instability (Grifferty et al.). Additionally, insufficient healthcare access further weakens the vulnerability of low-income populations. Many individuals in these communities lack health insurance or rely on underfunded public healthcare systems, leading to limited access to medical care and health systems (Wolf et al.). As climate change increases, environmental hazards such as air pollution and heat waves also increase, and residents of these communities are more likely to suffer from untreated respiratory diseases, cardiovascular conditions, and heat-related illnesses (*US EPA*). Exposure to high levels of air pollution has been linked to increased rates of anxiety and depression due to neuroinflammation and oxidative stress in the brain, which can impair cognitive function and alter emotions (Smirnova et al.). In turn, Low SES communities are disproportionately impacted by climate change due to their many disadvantages. Poverty makes it harder for low-income individuals to recover from climate disasters, while limited access to healthcare means they often struggle to get the mental and physical help they need. Further, poor housing conditions increase their exposure to hazardous environments, making these populations even more vulnerable.

Low-income communities face a disproportionate burden when it comes to climate hazards, as they are more likely to live in areas prone to extreme weather events, pollution, and incompetent infrastructure. Studies show that 40% of individuals exposed to flood risks globally come from low-income backgrounds (McDermott). These individuals often live in poorly insulated housing with outdated heating and cooling systems, which increases their vulnerability to extreme temperatures (*US EPA*). According to the Environmental Protection Agency (EPA), low-income neighborhoods tend to have less climate-adaptable infrastructures, making them more prone to damage from hurricanes, floods, and other natural disasters (OAR US EPA). Additionally, these communities face higher exposure to air pollution, with research showing that individuals from low socioeconomic backgrounds are more likely to live in areas with more dangerous levels of particulate matter and pollutants (Ballard Brief). Approximately, 70% of the world's poor population depends on natural resources for their livelihoods, making them more susceptible to climate-induced disruptions in agriculture, forestry, livestock, and fisheries (Mishra). In the United States alone, about 46 million people live in rural areas where healthcare access may be limited, increasing their vulnerability to climate-related health collisions (OA US

EPA). On top of that, Black and African American individuals are 34% more likely to live in areas with the highest projected increases in childhood asthma diagnoses and 40% more likely to reside in regions with the highest projected extreme temperature-related deaths with 2°C of global warming (Mishra). Furthermore, extreme climate events have displaced an average of 21.5 million people annually over the past decade, with 90% of these individuals coming from low-income developing countries and small island states (Mishra).

Studies show that lower-income households recover at a much slower rate after extreme weather events due to limited savings, lower insurance coverage, and lack of government assistance (National Low Income Housing Coalition). Poor living conditions further compound the risks posed by climate change. Many low-income individuals reside in standard housing at best with poor insulation, inadequate heating or cooling systems, and problems with the overall infrastructure that can make them more at risk of weather damage (Lucy Family Institute for Data & Society). Urban heat islands, areas with excessive concrete, fewer trees, and limited green spaces disproportionately impact low-income neighborhoods, exposing residents of the area to higher temperatures that can lead to heat exhaustion and dehydration, especially in young children and elderly individuals (Gochfeld and Burger). This prolonged exposure to extreme temperatures not only aggravates existing health conditions but also limits opportunities for relief as individuals may lack proper cooling units and struggle to pay the high energy costs that come with the amenities. Along with heat exhaustion, these communities are also often located near industrial zones or highways, resulting in chronic exposure to air pollution, which can increase the rate of asthma and other respiratory diseases (Gochfeld and Burger). These unfortunate realities prove that without proper resources, individuals in these areas are more likely to suffer from heat-related illnesses, further supporting the connection between low-income populations and climate change.

In addition to the physical damage from extreme weather events, low-income communities also face severe health risks due to improper sanitation and water quality (Wolf et al.). A review of drinking water, sanitation, and hygiene interventions highlights that poor sanitation and lack of clean water contribute significantly to the spread of illnesses like diarrhea, particularly in children of low-income areas (Wolf et al.). The combined stress of climate-related disasters and inadequate sanitation significantly compounds health risks in low-income communities, as floods and hurricanes often damage already fragile water and sanitation infrastructure, increasing risks of diseases (Wolf et al.). Finally, low-income communities are particularly vulnerable to climate hazards, as they live in high-risk areas with poor infrastructure and face heightened risks due to poor sanitation and limited resources.

Low socioeconomic (SES) communities face a disproportionate burden from climate-induced stressors, including extreme weather events and environmental pollution, which magnify existing health disparities. Research indicates that individuals in these communities are at a greater risk for mental health issues, partly because of inadequate access to mental health services and resources (Giudice et al.). Financial barriers make therapy and psychiatric care unaffordable, while shortages of mental health professionals in low SES areas are unfavorable

above all (Smirnova et al.). Therapy and physical care can be prohibitively expensive, with the average cost of a single therapy session ranging from \$100 to \$200 per session without insurance (Kuo and Tucci). Psychiatric care, including medication management, can be even more costly, with the initial consultations often exceeding \$300 to \$500, not including follow-up appointments which can range between \$100 and \$200 (Kuo and Tucci). For these low-income brackets, these costs are simply unaffordable. Additionally, insurance coverage for mental health services can be another major barrier for these communities. While the Affordable Health Care Act (ACA) and the Mental Health Parity and Addiction Equity Act (MHPAEA) aimed to improve coverage, many holes are yet to be filled (Frank et al.). In 2021, 10.2% of Americans remain uninsured, with the rate disproportionately higher for low-income individuals and communities of color (Frank et al.). Even those families with insurance may face high deductibles or copays that make treatment inaccessible (Frank et al.). In employer-sponsored insurance plans, the average deductible is \$1,669, a cost that many low-income individuals can not afford (Frank et al.). The inaccessibility of mental health care is further intensified by environmental stressors, which disproportionately impact low-income communities and contribute to worsening mental health conditions. The lack of awareness about available services exacerbates this issue, as residents may not know where to seek help (Giudice et al.). The combination of environmental stressors, limited government assistance, and pre-existing healthcare disparities creates a pattern of depreciation, significantly increasing the risk of mental health disorders such as anxiety, depression, and PTSD in low-SES communities. This amplifying effect makes these populations feel targeted toward mental health downfall, highlighting the urgent need for interventions and support.

Strategies for Immediate and Ongoing Mental Health Support

Despite the clear evidence of climate change's disproportionate impact on the mental health of low SES communities, there remains a significant lack of interventions. One major challenge is the long-lasting timeline required for mental health treatment, particularly when addressing both acute and chronic stressors. Acute stressors, such as single traumatic events like a natural disaster, can lead to immediate psychological distress, including PTSD anxiety, and depression (Feriante and Sharma). While some individuals may recover with short-term interventions, others require long-term interventions, long-term therapy, and support may be necessary for these other individuals to process the trauma (Feriante and Sharma). On the other hand, chronic stressors, such as prolonged exposure to poor-living conditions, financial instability, and environmental shame, gradually weaken mental health over time, often compelling the need for extensive treatment (Wang et al.). The collective effect of these stressors means that low-income individuals experience persistent mental health struggles, yet current interventions are short-term and reactive rather than preventive and supportive.

In the aftermath of extreme weather events such as hurricanes or wildfires, federal and state aid programs often prioritize infrastructure repair over individual mental health services, leaving many survivors without the necessary mental support that they need (Princeton

University et al.). Research indicates that only 1.2% of global health funding is dedicated to mental health services for climate-affected populations (Zhang et al.). Furthermore, communities with lower SES receive disproportionately fewer mental health services, with approximately 55% of low-income individuals reporting unsatisfied mental health needs (Zhang et al.). Moreover, rigid barriers, such as complex application processes, strict eligibility requirements, and long wait times, prevent many low-income individuals from receiving aid (OA US EPA). Studies indicate that individuals in lower-income brackets are less likely to apply for government relief due to a lack of awareness, distrust in institutions, or difficulty navigating the technical aspects (Mishra). Government-funded mental health programs are often overwhelmed, forcing individuals to endure long waitlists or receive quick and inefficient treatment (SAMHSA, *Greater Impact: How Disasters Affect People of Low Socioeconomic Status*). Without targeted efficient interventions, such as increased funding for mental health services, improved financial aid, and more effective application processes, the psychological toll on these communities will only worsen (SAMHSA, *Greater Impact: How Disasters Affect People of Low Socioeconomic Status*).

To protect vulnerable populations, policymakers must prioritize climate-durable planning in at-risk communities. Various sectors can play a role in addressing these disparities. Community-based mental health support systems, government initiatives, hospital/healthcare initiatives, and nonprofit organizations, all have significant potential to provide critical assistance. A stark example of climate-related mental health struggles was demonstrated when California's wildfires affected thousands of communities. The 2018 Campfire, the deadliest wildfire in the state's history, displaced over 50,000 people, a significant proportion was low-income (Hamideh et al.). Studies found that years following the disaster, nearly 40% of the survivors experienced PTSD, anxiety, and depression (Hamideh et al.). However, many lacked access to adequate mental health support, due to financial constraints and the destruction of local healthcare buildings (Hamideh et al.). Community-based mental health services, such as local support groups and peer counseling, can help individuals cope with climate-related stressors by providing accessible low-cost care (WHO, *Why Mental Health Is a Priority for Action on Climate Change*). Additionally, Medicaid expansion to cover climate-related mental health services is crucial, as approximately 74 million low-income families rely on Medicaid for access to healthcare (Carrie E. and Sommers). Research shows certain states that expanded Medicaid saw a 28% increase in mental health treatment rates among low-income populations (SAMHSA, *Greater Impact: How Disasters Affect People of Low Socioeconomic Status*). By implementing climate-resilient planning, expanding community-based mental health services, and increasing Medicaid coverage, policymakers and organizations can help mitigate the mental health impacts of climate tragedies, particularly for low-income populations who remain more vulnerable in these situations.

Expanding Medicaid to cover climate-related trauma interventions and investing in mobile mental health units could help address the problems that arise and ensure long-term mental health recovery for the affected populations. Mobile mental health units, which can

function as on-the-go clinics, have proven to be effective in delivering care to underserved areas, especially in the aftermath of disasters (Yu et al.). These units are equipped with licensed mental health professionals who provide counseling, crisis intervention, and psychiatric care directly in communities that lack easy access to traditional healthcare facilities (Yu et al.). At the peak of the Camp Fires, organizations such as Hope on Wheels Mobile Crisis Unit had stationed counselors in disaster zones, reducing response times and increasing access to care for displaced residents (Hamideh et al.). Reports show that mobile units led to a 32% decrease in emergency psychiatric hospitalizations and improved overall mental health outcomes in affected populations (Yu et al.). Integrating mobile mental health units into disaster response not only provides immediate psychological support but also reduces the long-term strain on emergency healthcare systems. Policymakers should prioritize increased funding and infrastructure for mobile mental health programs, ensuring they are always implemented into natural disaster aftermath plans. While expanding these programs requires financial funding, the cost is significantly lower than the long-term economic burdens of untreated mental health conditions (SAMHSA, *Greater Impact: How Disasters Affect People of Low Socioeconomic Status*). Studies show that every \$1 invested in mental health services yields a \$4 return in improved overall health and productivity (WHO). Conversely, the economic impact of unaddressed post-disaster mental health conditions including increased hospitalizations, lost work productivity, and higher rates of homelessness can lead to billions in costs (WHO). For example, the U.S. spent over \$200 billion on disaster recovery after Hurricane Katrina. Yet, mental health services were severely underfunded, leading many survivors to struggle with PTSD and anxiety for years to come (Princeton University et al.). By contrast, implementing mobile mental health units would require a fraction of the cost while significantly improving outcomes for affected communities (Yu et al.). Investing in mental health services alongside physical treatment ensures that disaster recovery is truly inclusive, addressing both the psychological and structural tolls of climate-related disasters.

Besides ensuring mobile mental health units, expanding telehealth services as a whole can further enhance accessibility, allowing residents to receive continued virtual counseling even after the mobile mental health units leave the area (SAMHSA, *Greater Impact: How Disasters Affect People of Low Socioeconomic Status*). Mental health recovery, especially for trauma-related conditions, can be a long-term process requiring constant therapy sessions, medication management, and emotional support. Many conditions, such as PTSD, depression, and anxiety, often require months or even years of treatment to fully manage the symptoms that can arise (WHO, *Why Mental Health Is a Priority for Action on Climate Change*). Through virtual therapy sessions, patients can maintain frequent consistency with the same counselors, reducing the disruptions of their treatment plans (Yu et al.). Additionally, telehealth services eliminate transportation barriers, which is crucial for those in low-income and rural areas where mental health providers are more scarce (Yu et al.). Given their success, these units should be a central component of climate resilience planning, ensuring that mental health is taken priority as much as physical disaster relief efforts.

Over and above mobile health units, hospitals can play an important role in helping climate disaster survivors by implementing proactive measures that address both physical and mental needs. Trauma-related care training is a measure that hospitals can implement to respond to the mental health needs of individuals (SAMHSA, *TIP 57: Trauma-Informed Care in Behavioral Health Services* | *SAMHSA Publications and Digital Products*). This type of training can certify that healthcare professionals understand the remaining effects of trauma, so they strive to create an environment where survivors feel safer and better supported during their recovery. Traditionally, healthcare environments have been more focused on symptom management or quick treatment of physical injuries, often disregarding the mental or emotional needs of patients (Sui et al.). This can lead to a “one size fits all” approach that doesn’t account for the unique needs of individuals who may have experienced serious trauma. In these particular situations, patients may not feel comfortable or safe sharing their experiences, and may even avoid seeking help altogether due to the fear of being misunderstood or re-traumatized (Sui et al.). In addition to mental health training in hospitals, the establishment of emergency room mental health groups can significantly improve the fallout for survivors (SAMHSA, *TIP 57: Trauma-Informed Care in Behavioral Health Services* | *SAMHSA Publications and Digital Products*). These teams can include psychologists, counselors, and social workers who can permanently be in ERs allowing immediate care for these individuals. Making this resource available early on for these individuals can halt further repercussions in the future. By incorporating climate-related mental health programs, hospitals not only support the immediate needs of natural disaster survivors but also help prevent long-term recovery, helping low-SES communities and individuals recover quicker and more efficiently.

Furthermore, community-based mental health programs can also close the gap in service accessibility. One example of a community initiative is the *Crisis Counseling Assistance Training Program (CCP)*, which provides federal funding to local agencies that offer crisis counseling and support services to those affected by disasters (CCP). This program has been especially beneficial towards low SES communities by hiring local counselors who understand the unique challenges that these individuals have faced (CCP). While these programs are effective, they are not widespread enough to reach the large number of individuals who need them, especially in rural and underserved areas (CCP). The widespread execution of these programs remains essential to ensuring that vulnerable communities receive the mental health support and resources they need to cope with climate-related disputes. Expanding these programs to reach more communities, with additional funding and coordination with local and federal governments could help close this gap by making mental health care more immediate, sustained, and accessible. In many cases, disaster relief efforts focus on short-term crisis response, yet mental health struggles persist long after the initial event. Providing consistent and long-term support such as follow-up counseling, trauma-informed therapy, and community outreach ensures that individuals do not face long-term problems without the care they need (Feriante and Sharma). Certifying that programs like the CCP receive continued funding and expansion efforts

can help establish a more consistent long-term mental health response in disaster aftermath plans, preventing vulnerable populations from falling through the cracks.

Nonprofit organizations and NGOs play a crucial role in providing mental health crisis support for individuals affected by climate disasters. For example, the National Alliance on Mental Illness (NAMI) provides resources and support for those experiencing mental health crises during natural climate-related disasters, with studies showing that mental health support programs can reduce the creation of severe conditions by 30-50% in affected populations (NAMI). As mentioned before, these programs are very small-scale and underfunded, which is a colossal barrier in place, preventing these programs from reaching the populations they need (NAMI).

However, it is important to recognize that the mental health struggles of low-SES communities due to climate change are not limited to large-scale disasters. Even in the absence of hurricanes, wildfires, or floods, the ongoing effects of climate change such as rising temperatures, poor air quality, and extended droughts contribute to chronic mental illnesses (Zhang et al.). Studies have linked extreme heat exposure to increased rates of aggression and suicide, particularly in low-income areas where residents may lack heating and cooling systems. The mental burden of living in areas that are prone to climate-related risks, often without the financial drive to relocate, creates consistent stress and amplifies mental health conditions (Hayes et al.). Finally, addressing the mental health needs of low SES communities in the climax of climate disasters requires many interventions, including the expansion of community-based programs, increasing funding for Medicaid and mental health services, and the implementation of mobile units and trauma care in ERs, to ensure that vulnerable populations receive the timely and effective support they deserve and need.

Conclusion

Climate change has wide-ranging consequences, particularly for low SES communities that face both immediate and long-term mental health challenges. Rising temperatures, natural disasters, and environmental hazards amplify stress, anxiety, and depression among vulnerable populations, who often lack the financial and healthcare resources to face these crises. However, low SES communities will not be the only communities disproportionately impacted by climate change, but also other marginalized communities, including racial and underserved minorities, immigrants, and individuals with disabilities, experience heightened risks due to inequalities from all angles. Without specific interventions, these risks and disparities will continue to grow, leaving millions of individuals without the mental health support they need (Hayes et al.). To address climate-related mental health disparities, policymakers must prioritize worthwhile and easily accessible interventions. Some of these interventions can include, expanding community-based mental health services, increasing Medicaid coverage for climate-related psychological conditions, and implementing more mobile mental health units into disaster response plans, which are essential to providing consistent care in times of a natural crisis (SAMHSA, Greater Impact: How Disasters Affect People of Low Socioeconomic Status).

However, long-term care strategies are also necessary to address the continuous mental health impacts of climate change. Sustainable investments in mental health infrastructure and preventive care can ensure long-term benefits for vulnerable populations. Bold measures such as trauma-informed care in hospitals can stop long-term mental harm. Also, guaranteeing that federal and state disaster relief programs allocate sufficient funding for mental health support, rather than solely focusing on infrastructure repair, can create a more well-rounded approach (Feriante and Sharma). Building climate resilience in vulnerable communities requires an encompassing long-term plan that involves mental health care for broader disaster response and recovery efforts. Expanding programs like the Crisis Counseling Assistance Training Program (CCP) and growing non-profit initiatives can fill the barriers, ensuring mental health support is not just available in the peak of disaster but remains a consistent line of support. To develop an adaptable future in healthcare, individuals must recognize that mental health recovery is just as essential as physical recovery. As climate change continues to grow the frequency and severity of extreme weather events communities will face both physical and psychological long-term trauma. Ensuring that vulnerable communities receive the mental health support they need to help them rebuild and adapt to the challenges that may lie ahead.

Works Cited

- Ballard Brief. "The Harmful Effects of Living in Brick Kiln Communities in the South Asia Region." *Ballard Brief*,
<https://ballardbrief.byu.edu/issue-briefs/the-harmful-effects-of-living-in-brick-kiln-communities-in-the-south-asia-region>. Accessed 2 Mar. 2025.
- Bienvenu, Germain. *Research Guides: Louisiana Hurricanes: Impact on Communities*.
<https://guides.lib.lsu.edu/Hurricanes/KatrinaCommunities>. Accessed 1 Mar. 2025.
- Carrie E., Fry, and Benjamin D. Sommers. "Effect of Medicaid Expansion on Health Insurance Coverage and Access to Care Among Adults With Depression." *Psychiatric Services (Washington, D.C.)*, vol. 69, no. 11, Nov. 2018, pp. 1146–52. *PubMed Central*,
<https://doi.org/10.1176/appi.ps.201800181>.
- CCP. *Crisis Counseling Assistance and Training Program (CCP)*. 7 Mar. 2024,
<https://www.samhsa.gov/technical-assistance/dtac/ccp>.
- Feriante, Joshua, and Naveen P. Sharma. "Acute and Chronic Mental Health Trauma." *StatPearls*, StatPearls Publishing, 2025. *PubMed*,
<http://www.ncbi.nlm.nih.gov/books/NBK594231/>.
- Frank, Richard G., et al. "Behavioral Health Parity and the Affordable Care Act." *Journal of Social Work in Disability & Rehabilitation*, vol. 13, no. 1–2, 2014, pp. 31–43. *PubMed*,
<https://doi.org/10.1080/1536710X.2013.870512>.
- Giudice, Linda C., et al. "Climate Change, Women's Health, and the Role of Obstetricians and Gynecologists in Leadership." *International Journal of Gynaecology and Obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics*, vol. 155, no. 3, Dec. 2021, pp. 345–56. *PubMed*, <https://doi.org/10.1002/ijgo.13958>.
- Gochfeld, Michael, and Joanna Burger. "Disproportionate Exposures in Environmental Justice and Other Populations: The Importance of Outliers." *American Journal of Public Health*, vol. 101, no. Suppl 1, Dec. 2011, pp. S53–63. *PubMed Central*,
<https://doi.org/10.2105/AJPH.2011.300121>.
- Grifferty, Grace, et al. "Vulnerabilities to and the Socioeconomic and Psychosocial Impacts of the Leishmaniasis: A Review." *Research and Reports in Tropical Medicine*, vol. 12, 2021, pp. 135–51. *PubMed*, <https://doi.org/10.2147/RRTM.S278138>.
- Hamideh, Sara, et al. "Wildfire Impacts on Education and Healthcare: Paradise, California, after the Camp Fire." *Natural Hazards (Dordrecht, Netherlands)*, vol. 111, no. 1, 2022, pp. 353–87. *PubMed Central*, <https://doi.org/10.1007/s11069-021-05057-1>.
- Hayes, Katie, et al. "Climate Change and Mental Health: Risks, Impacts and Priority Actions." *International Journal of Mental Health Systems*, vol. 12, no. 1, June 2018, p. 28. *BioMed Central*, <https://doi.org/10.1186/s13033-018-0210-6>.
- Impacts of Climate Change | US EPA*.
<https://www.epa.gov/climatechange-science/impacts-climate-change>. Accessed 1 Mar. 2025.
- Kuo, Dick C., and Veronica Tucci. "The Hidden Costs of Behavioral and Psychiatric

- Emergencies.” *Emergency Medicine Clinics of North America*, vol. 33, no. 4, Nov. 2015, pp. xvii–xviii. *PubMed*, <https://doi.org/10.1016/j.emc.2015.09.001>.
- Lucy Family Institute for Data & Society. *As Temperatures Rise, Research Points the Way to Lower Energy Costs, Better Living Conditions for Low-Income Households* - Lucy Family Institute for Data & Society. 6 Jan. 2025, <https://lucyinstitute.nd.edu/news-events/2025/01/06/as-temperatures-rise-research-points-the-way-to-lower-energy-costs-better-living-conditions-for-low-income-households/>.
- McDermott, Thomas K. J. “Global Exposure to Flood Risk and Poverty.” *Nature Communications*, vol. 13, June 2022, p. 3529. *PubMed Central*, <https://doi.org/10.1038/s41467-022-30725-6>.
- Mishra, Sushree. “Low-Income Communities Bear the Brunt of Climate Change.” *Earth.Org*, 9 Oct. 2023, <https://earth.org/climate-changes-unequal-burden-why-do-low-income-communities-bear-the-brunt/>.
- NAMI. “{OG: Title}.” *NAMI*, 30 Aug. 2017, <https://www.nami.org/Blogs/NAMI-Blog/August-2017/Helping-Others-Through-Catastrophic-Stress>.
- National Low Income Housing Coalition. *Study: Low-Income and Latino Households Face Disaster Recovery Challenges* | National Low Income Housing Coalition. 24 Feb. 2025, <https://nlihc.org/resource/study-low-income-and-latino-households-face-disaster-recovery-challenges>.
- NOAA. *Climate Change Impacts* | National Oceanic and Atmospheric Administration. <https://www.noaa.gov/education/resource-collections/climate/climate-change-impacts>. Accessed 19 Feb. 2025.
- Padhy, Susanta Kumar, et al. “Mental Health Effects of Climate Change.” *Indian Journal of Occupational and Environmental Medicine*, vol. 19, no. 1, 2015, pp. 3–7. *PubMed Central*, <https://doi.org/10.4103/0019-5278.156997>.
- Princeton University, Ushma Patel on, et al. *Hurricane Katrina Survivors Struggle with Mental Health Years Later, Study Says*. <https://www.princeton.edu/news/2012/01/24/hurricane-katrina-survivors-struggle-mental-health-years-later-study-says>. Accessed 1 Mar. 2025.
- Ramadan, Abdullah Mohammed Hassan, and Ahmed G. Ataallah. “Are Climate Change and Mental Health Correlated?” *General Psychiatry*, vol. 34, no. 6, 2021, p. e100648. *PubMed*, <https://doi.org/10.1136/gpsych-2021-100648>.
- SAMHSA. *Greater Impact: How Disasters Affect People of Low Socioeconomic Status*. ---. *TIP 57: Trauma-Informed Care in Behavioral Health Services* | SAMHSA Publications and Digital Products. <https://library.samhsa.gov/product/tip-57-trauma-informed-care-behavioral-health-services/sma14-4816>. Accessed 9 Mar. 2025.
- Smirnova, Natalia, et al. “Climate Change, Air Quality, and Pulmonary Health Disparities.”

- Clinics in Chest Medicine*, vol. 44, no. 3, Sept. 2023, pp. 489–99. *PubMed*, <https://doi.org/10.1016/j.ccm.2023.03.005>.
- Sui, Tiffany Y., et al. “The Impact of Physical Environments on Outpatient Mental Health Recovery: A Design-Oriented Qualitative Study of Patient Perspectives.” *PLOS ONE*, vol. 18, no. 4, Apr. 2023, p. e0283962. *PubMed Central*, <https://doi.org/10.1371/journal.pone.0283962>.
- US EPA, OA. *EPA Report Shows Disproportionate Impacts of Climate Change on Socially Vulnerable Populations in the United States*. 2 Sept. 2021, <https://www.epa.gov/newsreleases/epa-report-shows-disproportionate-impacts-climate-change-socially-vulnerable>.
- US EPA, OAR. *Impacts of Climate Change*. 16 Apr. 2021, <https://www.epa.gov/climatechange-science/impacts-climate-change>.
- USDA. *Economically Disadvantaged Communities | USDA Climate Hubs*. <https://www.climatehubs.usda.gov/hubs/northwest/topic/economically-disadvantaged-communities>. Accessed 19 Feb. 2025.
- Wang, Xiang, et al. “Acute Impacts of Extreme Temperature Exposure on Emergency Room Admissions Related to Mental and Behavior Disorders in Toronto, Canada.” *Journal of Affective Disorders*, vol. 155, Feb. 2014, pp. 154–61. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.jad.2013.10.042>.
- WHO. *Investing in Treatment for Depression and Anxiety Leads to Fourfold Return*. <https://www.who.int/news/item/13-04-2016-investing-in-treatment-for-depression-and-anxiety-leads-to-fourfold-return>. Accessed 16 Mar. 2025.
- WHO. *Why Mental Health Is a Priority for Action on Climate Change*. <https://www.who.int/news/item/03-06-2022-why-mental-health-is-a-priority-for-action-on-climate-change>. Accessed 9 Mar. 2025.
- Wolf, Jennyfer, et al. “Effectiveness of Interventions to Improve Drinking Water, Sanitation, and Handwashing with Soap on Risk of Diarrhoeal Disease in Children in Low-Income and Middle-Income Settings: A Systematic Review and Meta-Analysis.” *Lancet (London, England)*, vol. 400, no. 10345, July 2022, pp. 48–59. *PubMed*, [https://doi.org/10.1016/S0140-6736\(22\)00937-0](https://doi.org/10.1016/S0140-6736(22)00937-0).
- Yu, Stephanie W. Y., et al. “The Scope and Impact of Mobile Health Clinics in the United States: A Literature Review.” *International Journal for Equity in Health*, vol. 16, Oct. 2017, p. 178. *PubMed Central*, <https://doi.org/10.1186/s12939-017-0671-2>.
- Zhang, Shuo, et al. “Unequal Effects of Climate Change and Pre-Existing Inequalities on the Mental Health of Global Populations.” *BJPsych Bulletin*, vol. 45, no. 4, pp. 230–34. *PubMed Central*, <https://doi.org/10.1192/bjb.2021.26>.

History, Biology, and Advances in the Treatment of Alzheimer's Disease By Ethan Lao

Abstract

Alzheimer's Disease (AD) is the leading cause of dementia-related death, most commonly seen as brain atrophy that leads to memory loss and cognitive decline. Recognized by ancient physicians, it was formally diagnosed as a disease by Dr. Alois Alzheimer in 1906; he was also the first to recognize amyloid plaques and neurofibrillary tangles as key pathological features.

The tau and amyloid cascade hypothesis are central to modern AD research, with other factors such as cholesterol dysregulation, inflammation, and oxidative stress also being studied in their role in progressing AD. Genetic causes, though they remain few and between, are also investigated to ensure a complete view of the causes of AD.

Though no cure currently exists, many treatments for symptoms do, such as cholinesterase inhibitors and other drugs that target specific proteins, such as one that targets amyloid proteins. Ideas and treatments that may lead to cures include the use of CRISPR-Cas9 technology, as well as an experimental vaccine that may be used to prevent future generations from inheriting this disease.

Keywords

Alzheimer's Disease, Dementia, Amyloid plaques, Neurofibrillary tangles, Cognitive impairment, Neuropathology, Tau theory, Amyloid cascade theory

1. History

1.1 Ancient Views on Dementia

There is a very long history of Alzheimer's, beginning with the Ancient Egyptians and Greeks. Evidence shows that the Ancient Egyptians in 2000 B.C. knew that memory and mental function decline with age. In 570 to 495 B.C., the ancient Greek philosopher Pythagoras classified life into 6 stages: infancy, adolescence, adulthood, middle age, senescence, and old age. Pythagoras describes how it was expected in the stages of senescence and old age (70 to 80 years old) for mental functions to deteriorate and become mentally retarded. In 44 B.C., the ancient Roman philosopher Marcus Tullius Cicero wrote that dementia doesn't have to be caused by old age, something that hadn't been considered previously.⁹

1.2 Evolution of Dementia Understanding

In the 2nd century AD, a doctor in Turkey by the name of Aretheus described dementia as a form of delirium. He makes the key distinction that, while delirium is reversible, dementia is not, despite them both being disorders of cognitive function. In 1608, Shakespeare displayed signs of Alzheimer's Disease through the character Lear in his play *King Lear*. In 1672, the doctor Thomas Willis was the first to coin the term "neurology" as well as describing dementia in his book *De Anima Brutorum*, or, in English, *Of the Souls of Beasts*. In 1797, dementia was

accepted as a medical term and diagnosis by the French doctor Philippe Pinel. His disciple, Jean Etienne Dominique Esquirol, describes dementia as “disabilities that are shown in discernment, intellectual ability...and to lose joyfulness enjoyed.”⁹

1.3 Pioneers of Early Dementia Research

In 1894, Otto Ludwig Binswanger and Alois Alzheimer conducted studies on neurosyphilis, a cause of dementia. They report different forms of vascular dementia as a result of these studies, as well as coining the term “presenile dementia”. Alzheimer names it Binswanger’s Disease. In the late 1890s, a woman named Auguste Deter began to have trouble with memory, as well as experiencing behavioral changes. Though these are traditional symptoms of dementia, she is too young to have dementia, as she will only be turning 50 in 1900. As a result, from 1901 to 1902, Alzheimer consistently interviews Deter about her dementia-like symptoms. From these interviews, he learns that she is often disoriented, often calling items by the wrong name, and notices that her symptoms get worse at night. In 1906, Deter’s cognitive functions deteriorated almost completely due to a lack of treatment for her symptoms. She dies in April of that year.^{3, 9, 20}

1.4 The Naming of Alzheimer's Disease

After Deter’s death, Alzheimer requests her brain as well as her medical records for his studies. After his dissection, he notes that Deter’s brain has shrunk significantly in certain areas, as well as multiple abnormal deposits that he classifies as plaques and tangles. He concludes that Deter had a rare form of dementia that affects those under 65, and the plaques and tangles he discovers are related to her disease. In 1910, the German psychiatrist Emil Kraepelin classified dementia into two forms: senile and presenile dementia. The latter is named “Alzheimer’s disease” as this is the one discovered by Alois Alzheimer during his research with Auguste Deter.^{3, 9, 20}

1.5 Modern Era of Alzheimer's Research

In 1976, co-founder of the Alzheimer’s Association and neurologist Robert Katzman described Alzheimer’s disease as the most common cause of dementia, the fourth leading cause of death in the United States, and a public health issue. For the past 60 years prior to this, less than 150 scientific articles had been published on the research of Alzheimer’s, as most believed it was a rare form of dementia. After Katzman’s announcement, multiple organizations around the world were formed to fund research for cures and treatments for Alzheimer’s. More than 45,000 articles on Alzheimer’s have been published as well. Despite this, even in the present day, there is only medication to delay or reduce the effects of Alzheimer’s, and no cure has been discovered.^{9, 20}

2. Neurobiology of Alzheimer's

2.1 Amyloid Cascade Theory

The amyloid cascade theory revolves around the idea that the amyloid beta(A β) protein is what causes AD. In this case, the imbalance between the production and clearance of A β 42 and related A β peptides is what the amyloid cascade theory relies on. This begins when the amyloid precursor protein (APP) is mutated due to familial genetic mutations or naturally occurring random mutations. Due to these mutations, beta-amyloid (A β) proteins, specifically A β 42, begin to aggregate and oligomerize in the brain. This can lead to subtle effects on neuronal efficiency, as well as activating microglia and astrocytes, causing inflammation in the brain. Another result of this is beta-amyloid plaques forming. Due to increased production of A β proteins and inflammation, the clearance of these proteins is insufficient to keep the brain clear of A β . This causes plaques to form and continue growing as the process of AD continues.^{4, 8, 15, 17}

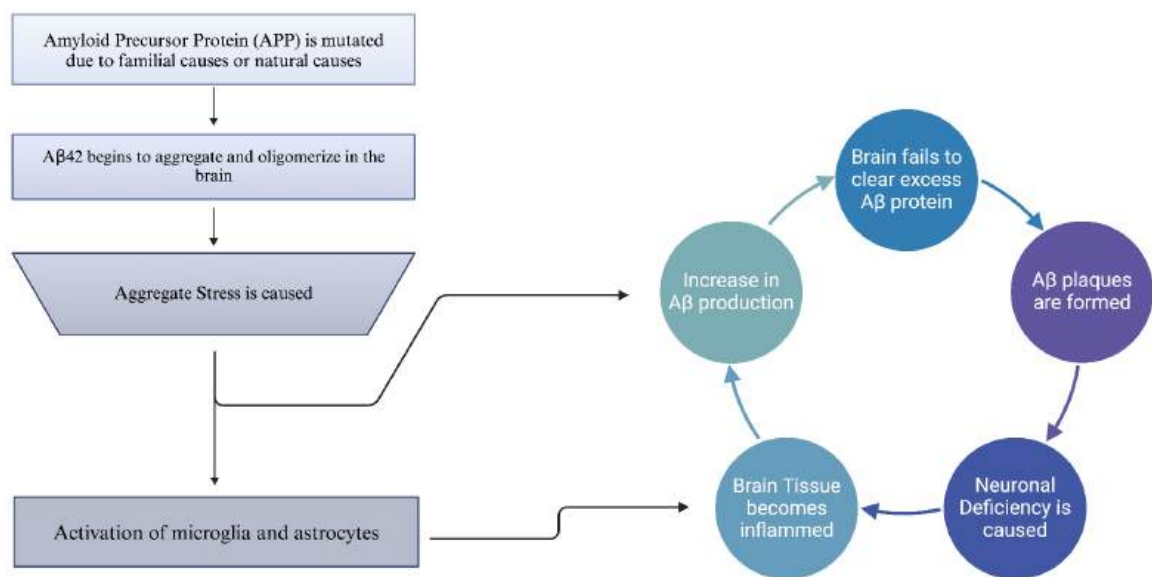


Fig 1: Flowchart of the Amyloid Cascade Theory

The validity of the amyloid cascade theory has fallen into doubt in recent years, as the experiment performed years ago that created the basis of this theory was not replicable. Not only that, but research has shown only familial AD is caused by aberrant amyloid expression. Late-onset AD (LOAD), also known as sporadic AD, is not triggered by aberrant amyloid expression, and neither is all early-onset AD caused by it. The amyloid cascade theory, therefore, is only a small portion of the causes of Alzheimer's disease.^{4, 13, 14, 15, 17}

2.2 Tau Theory

Another theory is the tau theory, where hyperphosphorylated tau is what causes AD. Normally, the protein tau is used in microtubules to support neurons in the brain. However, when these proteins are affected by genetic mutations, they become abnormally shaped and hyperphosphorylated, forming neurofibrillary tangles. Specifically, phosphorylated tau (p-tau) is the type found in these tangles, and they come in two forms: paired helical filaments with mixed

ribbons and straight filaments and non-fibrillated tau. These p-tau proteins are first found in distal dendrites before spreading to the soma. They target sensitive neurons, causing their spread to be seen across connected neural networks. The tau theory and amyloid cascade theory seem to be intertwined, as A β is known to increase abnormal phosphorylation of tau, while p-tau can trap APP in dendrites, causing more plaques to form in a vicious cycle.^{1, 15}

2.3 Cholesterol Levels

High cholesterol levels have recently been linked to an increased risk of AD. High cholesterol has been linked to more than a three-times increase in the chance of AD, because it can lead to an increase of plaques forming in the brain. The main reason for this has been attributed to high cholesterol causing altered cholesterol processing in brain cells, disrupting normal functioning. This leads to impaired neuronal communication, which makes it easier for cells to become more easily affected by plaques and tangles. Other research has found that oxidized cholesterol metabolites, formed when cholesterol is broken down, is particularly dangerous to brain cells. Although not a direct cause of AD, it does have a strong association with the disease.¹⁶

2.4 Inflammation

Chronic inflammation in the brain is a significant contributor to the progression of AD. Microglia are the immune cells in the brain, but they fight issues in the brain cells by causing inflammation. This normally has a protective function but when microglia is constantly activated by A β , it is detrimental to the brain. Microglia begin producing toxic byproducts and increased proinflammatory proteins that cause even more damage to the brain, such as reactive oxygen species, nitric oxide, and cytokines. Elevated levels of microglia are also found to stimulate the phosphorylation of tau, creating tangles. This may be a link between the two theories of AD.⁸

2.5 Oxidative Stress

Oxidative stress has a significant role in the process of AD. It is a result of an imbalance in the redox state of the brain. It involves the generation of excess reactive oxygen species (ROS) and the dysfunction of the antioxidant system. The mitochondrial electron transport chain creates byproducts in the form of hydrogen peroxide and superoxide radicals. When overproduced, these can result in tissue damage that can cause free radical reactions due to reacting with excess iron. Because the brain has a high demand for oxygen, it is a prime location for oxidative stress to occur. When these reactions occur, acidic compounds that are detrimental to brain tissue form, creating a significantly higher risk of AD forming.^{12, 20}

3. Progression of Alzheimer's

3.1 Phases of Alzheimer's

There are 3 main phases of the progression of Alzheimer's disease.

The first phase is when the afflicted patient is asymptomatic and acts normally. However,

signs of amyloid deposition in the brain still exist, despite the lack of change in behavior. Scans during this phase often show normal brain function despite the amyloid deposition.

The second phase is when the afflicted patient is symptomatically prodromal, or in other words, begins showing early signs of the disease. The patient may experience mild cognitive impairment or other symptoms due to an increasing amount of amyloid deposition. It has been noted that cognitive decline in those with Alzheimer's is at a higher rate compared to individuals of the same age without cognitive impairment. However, in those without Alzheimer's, the cognitive decline doesn't result in impaired functioning.

In the last phase, the patient's mental decline results in an inability to live normally due to cognitive impairment. The patient is officially diagnosed with dementia, and the body itself has gone through molecular changes. These changes are attributed to the neurodegenerative properties of Alzheimer's.^{4, 15, 18}

3.2 Genetics of Alzheimer's

Because Alzheimer's is genetic, many genes can affect or cause the disease. One of these is the gene Apolipoprotein E or APOE. It is involved in the cholesterol transport in the cerebral spinal fluid, as well as the binding and clearance of beta-amyloid in the brain. The gene consists of 3 alleles: $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$. The $\epsilon 4$ allele has the greatest risk of developing late-stage familial and sporadic Alzheimer's because it reduces cholesterol efflux from neuronal cells while also increasing beta-amyloid binding and deposition. This allele is located in 15% of the total population, but 40% of all Alzheimer's patients. On the other hand, the allele $\epsilon 2$ appears to have a protective property against Alzheimer's.^{4, 15, 16}

APOE is not the only gene to affect Alzheimer's. Many other genes are known to affect the affixation, processing, and transportation of beta-amyloid, such as ABCA7, clusterin (CLU), and sortilin-related receptor L (SORL1). Rare, heterozygous variants of the gene TREM2 are associated with a significant increase in the risk of Alzheimer's disease.¹⁵

Other research shows that microRNA (mRNA) possibly affects Alzheimer's disease by interrupting the metabolism of amino acids in the brain if it is improperly regulated. For example, one form of mRNA, miR-153, inhibits the expression of amyloid precursor proteins (APP). When levels of miR-153 are low, this drives increased APP expression in some Alzheimer's patients. Another form, miR-339-5p, has also been shown to be dysregulated in the brain of Alzheimer's patients.^{4, 15}

3.3 Forms of Alzheimer's Disease

A small percentage of Alzheimer's patients have developed the disease at an early age. Studies done on these patients have found that early-onset Alzheimer's (EOAD) is a result of genetic causes. The genes presenilin 1(PSEN1) and presenilin 2(PSEN2) are components of the gamma(γ)-secretase complex. γ -secretase is a complex needed for the production of amyloid beta(A β). If amyloid beta 40 or 42 has a mutation, this results in the protein having a higher processing rate through the amyloidogenic pathway, resulting in a build-up in the brain. When

amyloid beta isn't mutated, it is transferred to the cerebral spine fluid at a rate that ensures that the buildup doesn't form in the brain. However, when the amyloid beta is mutated and does buildup, beta-amyloid plaques form at an increased rate. This process only occurs in a small number of cases.⁴

Sporadic late-onset AD (LOAD) is classified by symptoms showing past ages of 65. It is the most common form of Alzheimer's, and, although not the disease is not guaranteed by any specific gene, it has a higher risk of developing due to the genetic risk factor of APOE. However, more than 20 other genetic factors have been identified that can affect the chance of developing AD. Similar to EOAD, mutations in PSEN1 and PSEN2 have been identified in those with LOAD, but at a much lower percentage of patients. APOE ϵ 4 is also known to reduce the time of getting AD by a century for each APOE ϵ 4 allele present. Female carriers are also at a higher risk of getting LOAD from the ages 65-75 compared to male carriers. Some non-genetic risk factors for LOAD are traumatic brain injury, an unhealthy diet, lack of sleep, and smoking. Those who have had depression also may have a higher chance of developing LOAD in their lifetime, especially those with late-onset depression.¹⁴

4. Neuropathological Changes of Alzheimer's

There are 3 major forms of neuropathological changes in the brain related to Alzheimer's disease: beta-amyloid plaques, neurofibrillary tangles, and brain atrophy.

Beta-amyloid plaques, as previously described, are a result of buildup in the protein of amyloid beta in the brain. These plaques can be discovered in the brain through PET scans, post-mortem examinations, or clinical tests that record low amounts of A β 1-42 in the cerebral spinal fluid. The most widely accepted hypothesis of how these plaques form is the amyloid cascade hypothesis; this idea proposes that an imbalance between the production and clearance of amyloid beta causes Alzheimer's disease. Normally, A β is cleared from the central nervous system through vascular mechanisms, phagocytosis by microglia, enzymatic breakdown, or drainage in the brain's capillary basement membranes. These membranes also drain fluids and other solutes from the brain. When A β is deposited in these membranes during cerebral amyloid angiopathy, A β elimination and interstitial fluid drainage are impeded during Alzheimer's disease. This can cause slow changes in the concentration of A β in the brain, causing a cascade of neurological effects throughout the body, resulting in neuronal dysfunction, cell death, or dementia. Increased A β levels in the brain can also be caused by overproduction, such as in the cases of early-onset Alzheimer's, or reduced efficiency of clearance, such as in late-stage Alzheimer's.¹⁵

Neurofibrillary tangles are a result of increased amounts of intraneuronal fibrils that mostly consist of abnormal, hyperphosphorylated forms of the protein tau. Tau is usually associated with microtubules in neuron cells, as its role is to stabilize the microtubules in axons, providing support for the cell's internal transport system. Finding increased tau in the cerebrospinal fluid is not necessarily a sign of Alzheimer's disease, but is usually a sign of increased clinical disease severity. Phosphorylated tau (p-tau) is the protein in neurofibrillary

tangles and the amount of p tau is more directly related to Alzheimer's disease than the total amount of tau (t-tau). Hypophosphorylated tau is found in two parts of the brain. Firstly, it is found in neurofibrillary tangles of paired helical filaments with mixed ribbons and straight filaments. Secondly, non-fibrillized tau is located in the cytosol of the brain. In the brains of those affected by Alzheimer's disease, the levels of tau protein are 4 to 8 times more prevalent compared to control brains. Hypophosphorylated tau cannot make microtubules and also isolates normal tau (microtubule-associated protein 1 and 2 (MAP1 and MAP2)), resulting in the inhibition and disassembly of microtubules. As tau becomes more and more phosphorylated, it becomes insoluble and self-assembles into wide, loosely coiled paired helical filaments and more tightly coiled straight filaments, resulting in the neurofibrillary tangles forming. However, after the tangles form, the isolation of other tau is no longer possible; this is theorized to be a self-defense mechanism created by the neurons. Mouse models suggest that the spread of Alzheimer's disease-related neuronal dysfunction is through connected and vulnerable neurons in an anatomical network.¹⁵

Brain atrophy is a direct result of Alzheimer's disease. Parts of the brain shrink, resulting in reduced functionality throughout the body.^{9, 12} The rate and distribution of brain atrophy is consistent and predictable, allowing patterns of brain atrophy to be predictive of dementia and other cognitive diseases. These patterns can be studied using a vMRI, allowing the probability of Alzheimer's disease to be known through the presence of specific biomarkers. These biomarkers are as follows:

1. Increase of tau in the cerebral spinal fluid
2. Decrease of A β in the cerebral spinal fluid as it concentrates in the brain
3. Increase of amyloid deposition through PET scan
4. Brain atrophy (can be measured using MRI, fMRI, and FDG PET)
5. Cognitive impairment
6. Decreased clinical functioning and activities in daily living^{15, 18}

5. Cures and Treatment Options

As of when this was written, there are no cures for Alzheimer's. There are, however, many treatments designed to reduce the severity of and control the symptoms of Alzheimer's.

Some of these drugs, such as galantamine, benz galantamine, rivastigmine, and donepezil, are used to inhibit cholinesterase, resulting in the breakdown of acetylcholine being prevented. Acetylcholine is a chemical important for memory and thinking, and when it is broken down, memory and thinking capabilities are heavily reduced. These drugs lose their effectiveness over time because the brain stops producing as much acetylcholine throughout Alzheimer's. Switching drugs may not be more or less effective because they all have the same purpose, but patients may respond better to some drugs compared to others.

Lecanemab and donanemab are immunotherapy drugs designed for the early stages of Alzheimer's disease. These drugs target A β , the protein that causes plaques in the brain. Medical

studies and trials have shown that, while using these drugs, patients have a decreased rate of cognitive decline. This is a direct result of the reduction of amyloid in the brain.

Memantine is another drug that reduces the symptoms of Alzheimer's. It may allow patients to perform daily functions a little longer than normal. This drug regulates glutamate, which, when produced in excess amounts, results in cell death. This cell death can speed up the process of Alzheimer's.

5.1 Current Research on Cures

There is some research being done on possible future cures for Alzheimer's that have not been tested on humans but have shown promising results on mouse models.

One cure uses CRISPR/Cas9 technology to edit the mutated genes that lead to Alzheimer's disease. The difficulty of doing so currently limits the usage of this technology. Many professionals in the field remain doubtful that, with our current technology, CRISPR/Cas9 can completely eradicate Alzheimer's, but they still believe that it can significantly reduce the effects or causes of the disease. Mouse models testing the technology have demonstrated that CRISPR is effective at reducing amyloid production and removing genes that increase the risk of Alzheimer's is possible.^{6, 19}

There is also a vaccine being developed for the prevention of Alzheimer's. Mouse models testing this vaccine have shown that amyloid deposits in the brain were significantly reduced in size, with inflammatory biomarkers being reduced. The vaccine works by targeting cells with SAGP proteins. SAGP stands for senescence-associated glycoprotein, which is typically found to be overexpressed in inflamed brain cells. When applied, the vaccine is thought to clear the SAGP proteins. The protein is most found in microglia, which plays a role in the immune system while also causing inflammation. This inflammation damages the neurons, so by targeting SAGP, microglia is also reduced, causing more neurons to be saved. Mice treated with these vaccines have been shown to act more like normal mice compared to their placebo counterparts.^{2, 7, 11}

Mouse models have been used in the testing of treatments and possible cures for Alzheimer's disease. Using technologies like single-cell and single-nucleus RNA sequencing and spatial transcriptomics, research has shown that responses at the single-cell level can lead to differing susceptibility to A β toxicity of neurons. This allows researchers to switch mice genes with human genes, such as specific alleles that affect A β expression, to test for their effects on the mice. However, the breeding of these mice for specific existing alleles would also affect the expression of the genes when spliced in, making this a process that will need to be refined over time.⁵

6. Discussion

Overall, Alzheimer's disease is a long, complicated disease that stems from multiple causes. There is not one known cause of Alzheimer's, and research being done is working on narrowing down the factors for the development of the disease. Setbacks such as the controversy surrounding the falsification of results of beta-amyloid plaques have caused some discussion

over the other possible causes of the disease as well. Understanding the underlying risk factors, from lifestyle to genetics, and the interplay between current theories such as the amyloid cascade and phosphorylated tau, is central to discovering the treatment and cure for Alzheimer's.

7. Conclusion

As technology improves, so will the ability to detect, treat, and possibly even cure Alzheimer's. This disease is one that heavily affects the lives of patients and those around them. Each discovery strengthens our ability to fight against this terrible disease. As Alzheimer's is now a world health concern, the Alzheimer's Association has updated the NINCDS-ADRDA criteria for a stricter, more specific way to diagnose patients with the disease. Recent studies have given a chance at preventing Alzheimer's, but the goal currently is to create more potent and effective medications to lessen the symptoms of the disease and prolong the time that a patient can remain self-sufficient. This may come, in time, with the increasing knowledge of the causes behind Alzheimer's disease, and the technological capabilities to create more specific and stringent testing protocols to narrow down said causes. Designing a drug that is potent, reliable, and available to all is the key to treating patients with Alzheimer's and preventing more from getting the disease.

Works Cited

1. Arnsten, Amy F. T., et al. "Hypothesis: Tau Pathology Is an Initiating Factor in Sporadic Alzheimer's Disease." *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, vol. 17, no. 1, Jan. 2021, pp. 115–24. PubMed, <https://doi.org/10.1002/alz.12192>.
2. Association, American Heart. Novel Vaccine May Hold Key to Prevent or Reduce the Impact of Alzheimer's Disease. <https://medicalxpress.com/news/2023-07-vaccine-key-impact-alzheimer-disease.html>. Accessed 02 Feb. 2025.
3. Bondi, Mark W., et al. "Alzheimer's Disease: Past, Present, and Future." *Journal of the International Neuropsychological Society : JINS*, vol. 23, no. 9–10, Oct. 2017, pp. 818–31. PubMed Central, <https://doi.org/10.1017/S135561771700100X>.
4. Breijyeh, Zeinab, and Rafik Karaman. "Comprehensive Review on Alzheimer's Disease: Causes and Treatment." *Molecules*, vol. 25, no. 24, 24, Jan. 2020, p. 5789. www.mdpi.com, <https://doi.org/10.3390/molecules25245789>.
5. Chiu, Yuanpu, et al. "Genetically Engineered Mouse Models for Alzheimer Disease and Frontotemporal Dementia: New Insights from Single-Cell and Spatial Transcriptomics." *The American Journal of Pathology*, Dec. 2024, pp. S0002-9440(24)00447-4. PubMed, <https://doi.org/10.1016/j.ajpath.2024.11.006>.
6. De Plano, Laura M., et al. "Applications of CRISPR-Cas9 in Alzheimer's Disease and Related Disorders." *International Journal of Molecular Sciences*, vol. 23, no. 15, Aug. 2022, p. 8714. PubMed, <https://doi.org/10.3390/ijms23158714>.
7. Hovakimyan, Armine, et al. "mRNA Vaccine for Alzheimer's Disease: Pilot Study." *Vaccines*, vol. 12, no. 6, 6, June 2024, p. 659. www.mdpi.com, <https://doi.org/10.3390/vaccines12060659>.
8. Kinney, Jefferson W., et al. "Inflammation as a Central Mechanism in Alzheimer's Disease." *Alzheimer's & Dementia : Translational Research & Clinical Interventions*, vol. 4, Sept. 2018, pp. 575–90. PubMed Central, <https://doi.org/10.1016/j.trci.2018.06.014>.
9. Knopman, David S., et al. "A Brief History of 'Alzheimer Disease.'" *Neurology*, vol. 92, no. 22, May 2019, pp. 1053–59. PubMed Central, <https://doi.org/10.1212/WNL.00000000000007583>.
10. Kumar, Ashutosh, et al. "ADULT NEUROGENESIS IN HUMANS: A Review of Basic Concepts, History, Current Research, and Clinical Implications." *Innovations in Clinical Neuroscience*, vol. 16, no. 5–6, May 2019, pp. 30–37.
11. Lambracht-Washington, Doris, and Roger N. Rosenberg. "Advances in the Development of Vaccines for Alzheimer's Disease." *Discovery Medicine*, vol. 15, no. 84, May 2013, pp. 319–26.
12. Machado-Oliveira, Gisela, et al. "Cell Senescence, Multiple Organelle Dysfunction and Atherosclerosis." *Cells*, vol. 9, no. 10, 10, Oct. 2020, p. 2146. www.mdpi.com,

<https://doi.org/10.3390/cells9102146>.

13. Morris, Gary P., et al. "Inconsistencies and Controversies Surrounding the Amyloid Hypothesis of Alzheimer's Disease." *Acta Neuropathologica Communications*, vol. 2, Sept. 2014, p. 135. PubMed, <https://doi.org/10.1186/s40478-014-0135-5>.
14. Rabinovici, Gil D. "Late-Onset Alzheimer Disease." *Continuum : Lifelong Learning in Neurology*, vol. 25, no. 1, Feb. 2019, pp. 14–33. PubMed Central, <https://doi.org/10.1212/CON.0000000000000700>.
15. Raskin, Joel, et al. "Neurobiology of Alzheimer's Disease: Integrated Molecular, Physiological, Anatomical, Biomarker, and Cognitive Dimensions." *Current Alzheimer Research*, vol. 12, no. 8, Oct. 2015, pp. 712–22. PubMed Central, <https://doi.org/10.2174/1567205012666150701103107>.
16. Sáiz-Vazquez, Olalla, et al. "Cholesterol and Alzheimer's Disease Risk: A Meta-Meta-Analysis." *Brain Sciences*, vol. 10, no. 6, June 2020, p. 386. PubMed Central, <https://doi.org/10.3390/brainsci10060386>.
17. Selkoe, Dennis J., and John Hardy. "The Amyloid Hypothesis of Alzheimer's Disease at 25 Years." *EMBO Molecular Medicine*, vol. 8, no. 6, June 2016, pp. 595–608. PubMed Central, <https://doi.org/10.15252/emmm.201606210>.
18. Tahami Monfared, Amir Abbas, et al. "Alzheimer's Disease: Epidemiology and Clinical Progression." *Neurology and Therapy*, vol. 11, no. 2, June 2022, pp. 553–69. PubMed, <https://doi.org/10.1007/s40120-022-00338-8>.
19. Thapar, Nandita, et al. "Application of CRISPR/Cas9 in the Management of Alzheimer's Disease and Parkinson's Disease: A Review." *Annals of Medicine and Surgery*, vol. 86, no. 1, Nov. 2023, pp. 329–35. PubMed Central, <https://doi.org/10.1097/MS9.0000000000001500>.
20. Huang, Wen-Juan, et al. "Role of Oxidative Stress in Alzheimer's Disease (Review)." *Biomedical Reports*, vol. 4, no. 5, May 2016, pp. 519–22. www.spandidos-publications.com, <https://doi.org/10.3892/br.2016.630>.
21. Yang, Hyun Duk, et al. "History of Alzheimer's Disease." *Dementia and Neurocognitive Disorders*, vol. 15, no. 4, Dec. 2016, pp. 115–21. PubMed, <https://doi.org/10.12779/dnd.2016.15.4.115>.

The Influence of Color on the Human Mind:

An exploration on the Cognitive and Emotional Responses By Tehreem Arshad

Abstract

The complex relationship between the human brain and color is a multifaceted field of research that incorporates psychology, neuroscience and behavioral science. This paper discovers how exactly a differing variety of colors have the ability to influence moods, basic functions and psychological well being in regards to current academic journals and ongoing studies. Through analyzing how certain colors are able to cause reactions from certain regions of the brain, such as how the occipital lobe is responsible for visual processing or how the limbic system interconnects and works towards a final goal of providing humans with emotional reactions. With the goal of understanding the effects of different colors on the human mind it is necessary to extend that research and apply it to the basis of the world and our day-to-day lives. The application of this research extends to a wide range of fields including education, marketing, retail and therapeutics where color is strategically utilized to manipulate human behavior. Along with the neural groundwork of color perception there are also cultural and individual differences that are taken into account, to show the intricacy and great depths of this topic.

Introduction

Colors are foundational elements in the lives of most humans that have the ability to shape how we perceive this world. This can range from the empty black hues in the darkness under the bed as kids or the calming purple and blue shades of the night sky. As young children, individuals are exposed to a vast range of colors that all have common connotations associated with them. With certain colors causing fear and anxiety while others causing happiness and excitement. Within these past couple of years studies have shown that colors have a sizable impact within human behavior including emotions. This has caused an uprising in terms of interest in the ways that certain colors can be used to manipulate human behavior whether that be in marketing, art or therapy. However, even with the extensive information already known today in terms of the psychological effects of colors there still is a gap in the understanding of how different colors affect the mind on a deeper level.

Although many studies have researched and discovered the psychological impacts of color, the exact bodily processes that explain this phenomenon aren't fully understood. Understanding these connections between color and mental processes could give valuable insights into how colors shape emotional responses, memory retention, and decision-making. The importance of this research lies in its many applications. By going over the biological and psychological connections between color and emotion, the development of more effective strategies in many fields could occur, such as advertising, education, healthcare, and mental health treatment.

The Neural Mechanisms of Color Perception

The process of color perception all begins in the eyes, since that is where light is first received and transmitted to the brain. At first as the light comes into the eye, it begins to be processed by these special cells in the retina, known as the photoreceptors, that are hyper focused on converting light into electrical signals allowing for humans to see. The retina contains two primary photoreceptors, which are rods and cones. Rods have the responsibility of allowing humans to see in low amounts of light while cones are necessary for detecting light in the visible spectrum. Humans have three types of distinct color cones that detect red, blue and green light and there are about six million within the eye area, and most of them are packed into the fovea which is a small pit in the back of the eyes. From there, these signals travel through the optic nerve into some of the brain's regions of visual refinement which include the primary visual cortex, the V4 area and the inferotemporal cortex. In the back of the occipital lobe is where the primary visual cortex is found and it is the first to process visual information (Shapely & Hawken, 2002)lw. Afterwards, color signals go into the V4 area where it becomes distinguished from other hues and is compared to other colors. Once color signals have finally made their way out of the V4 area they proceed into the inferotemporal cortex, which allows for people to recognize elaborate patterns and make mental connections of these patterns to specific memories and feelings and this is known as the feeling of nostalgia. Some frequent connections based on personal experience made with certain colors include yellow for happiness, blue for sadness and red for anger but these can differ for individuals (Knez and Kers, 2000). Along with that, another part of the brain that plays an important role in the color perception process is the limbic system. The function of the limbic system is to regulate one's emotions such as fear, excitement or aggression (Elliot & Maier, 2014).

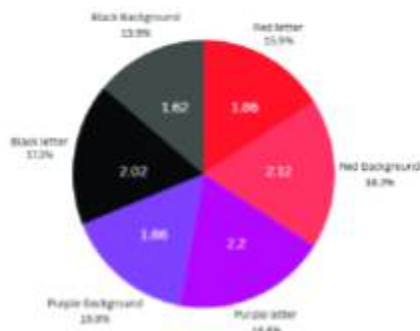
Overall, color perception is a vast and complicated process that involves many regions of the brain to work together. Whether that is from the retina's photoreceptors to the visual cortex, the V4 area, the inferotemporal cortex, the limbic system or the optic nerve. The human brain has the power to not only process a wide shade range but to correlate these colors to past events to cause an emotional response. Due to this beautiful process we are able to understand why colors can be so relevant to human behavior and why colors become conditioned to adhere to a certain behavior.

Cognitive Performance in Regards to Color

Color does way more than influence our emotions, it also plays an important role in cognitive performance. Many researchers have noticed that certain colors can increase or decrease the efficiency of different cognitive tasks, such as attention, memory, and problem solving abilities. These specific observations were not just individual results that relied on the person being studied, rather they are dependent on the way the brain comprehends peculiar visual stimuli and how our cognitive processes react to these stimuli.

One of the most well researched effects of color on cognitive performance is its impact on attention and focus. A variety of studies have shown that warm colors such as red or yellow

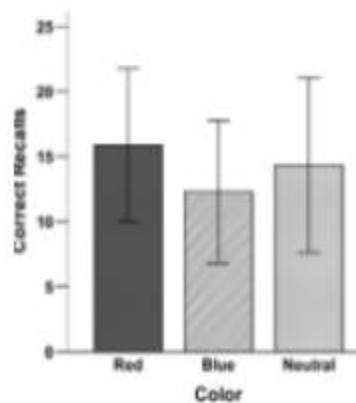
can increase attention and alertness, while cooler colors such as blue or green can have a calming effect that may reduce distractions (Elliot & Maier, 2014). For instance, research released by Kwallek, Lewis, and Robbins (1996) observed that the workers in environments with red or yellow backgrounds performed better on tasks that were heavy in terms of attention and vigilance, while blue or white environments led to better performance in tasks requiring people to use their creativity. This noted difference is thought to occur because of the physiological responses to these colors which for example would be, warm colors causing an increased arousal, which can lead to an improved performance on tasks centered on attention. Along with that color also takes part in memory retention and recall. Morrow, Leirer, and Finkelstein (2004) were all researchers that explored how color affects the recall of words and they discovered that subjects were able to remember colored words more easily than black-and-white words. Color usage in written materials, such as textbooks, flashcards or notes can make information more and easier to recall and is often a method used by many teachers to boost memorization of the content of a topic.



Roy, A., & Thattil, J. (2024). *EFFECTS OF WARM, COOL, AND NEUTRAL COLORS ON SHORT TERM MEMORY IN YOUNG CHILDREN*. Research Gate. Retrieved March 31, 2025, from https://www.researchgate.net/figure/Pie-Chart-Illustrates-the-mean-scores-of-different-colors-black-red-and-purple-on_fig1_38454173

The study suggested that certain colors can help link information to the color so that it stands out in memory later on. This event occurs because colors tend to be visual cues in our memory that make information more distinctive than other relevant memories. Moreover, the researcher Knez (2001) noted that the use of color in learning environments can improve both short-term and long-term recall as well. For example, when students were told to study in rooms with colored walls they had the ability to recall studied content better than those students who studied in neutral-colored rooms. His work demonstrates that color enhances both the significance of information and it increases the emotional association that goes along with it which leads to a stronger memory formation. However, the impact of color surpasses attention

and memory since it can also influence creativity and problem-solving. Some studies have indicated that different colors can either revive or discourage creative thinking. For example, the color red has been shown to increase focus and precision, making it perfect for attention heavy tasks with detail precision and organized thinking involved (Mehta & Zhu, 2009). Furthermore, blue is linked with magnified creativity, since it's often fused to more relaxed states of beings that encourage open-mindedness and innovative thinking. Mehta and Zhu (2009) were researchers who conducted an experiment in which participants were placed into either a red or blue room before they were asked to perform tasks with creativity involved. The subjects in the blue environment performed at a higher rate on tasks requiring creative thinking and problem-solving, while participants in the red environment showed improved performance on tasks that required attention to detail and accuracy.



Mehta, R., & Zhu, R. (2009). *Blue or Red? Exploring the Effect of Color on Cognitive Task Performances*. Science Advisor. Retrieved March 31, 2025, from <https://www.science.org/doi/10.1126/science.1169144>.

Mehta and Zhu argued that this may be because blue is already associated with a calm, relaxed state which is crucial to creative thinking, while red increases arousal, which is more beneficial for more focused tasks. The research on color and cognitive performance has more of an increased significance for how the design of learning and work environments is done. Such as, classrooms and offices being built with specific colors to either encourage focus and productivity or encourage creativity. Orange and yellow are both warm colors that have stimulating effects, could be ideal for cooperative spaces, while blue could be utilized in areas meant to promote deep thinking and problem-solving skills. Similarly, the color green is often associated with balance and calm, and could be efficient in spaces where the stress levels are meant to stay low such as libraries or study rooms. Moreover, red could be strategically used in

high-stress environments such as emergency rooms, where alertness and precision are absolute. Conversely, spaces built for relaxation, such as spa or therapy areas, might do better off from more somber tones like light blue or green, which have a pattern on reducing anxiety and promoting a calm state (Elliot & Maier, 2014). The effects of color on cognitive performance are solid, with evidence from many different research papers and experiments suggesting that color can influence attention, memory, creativity, and overall cognitive functioning.

The Emotional Aspect of Colors

In color psychology, which has been well researched, colors have a strong power to affect human emotions. The emotional reactions colors cause are based on cultural association, personal experience and physiological processes; they are not simply subjective experiences. The brain deals with colors in a way that stimulates particular emotional responses, and these responses may differ depending on personal experience, cultural background, even physiological reactions. Moreover the knowledge of this process can help people across many different fields including marketing, teaching, and health care, as well though color decisions can affect attitude, activity, and performance. Warm colors are usually linked with arousal and increased emotional reactions such as red, orange, and yellow. These hues are recognized to increase alertness, warmth, and activity level as well as boost energy levels. Red, specifically, is usually associated with strong feelings of hate, anger, passion, and risk (Elliot & Maier, 2014). Red is often utilized in marketing to draw attention and encourage activity for example, fast-food branding or sales signs. To discover the physiological reason, the color red is believed to stimulate the arousal systems of the brain, which raises heart rate and blood pressure (Wright & Kipp, 2011). On the other hand, while still warm, orange tends to stimulate inspiration, hope, and ingenuity. Many people view it as inviting and inspiring, and it has been connected with greater social interaction and communication (Birren 2013). Another warm color, yellow, is usually linked with joy, optimism, and mental clarity and is known to excite the mind and promote focus, which causes it to be a favorite selection in circumstances that demand mental sharpness such as study spaces or classrooms (Kwallek & Lewis, 1990). The mind usually finds relaxation and calmness in cool colors like blue, green, and purple. Blue is well known for its power to generate emotions of rest, peace, and trustworthiness (Elliot & Maier, 2014) which is why blue is often chosen for surroundings where focus and calm are needed, such as bedrooms, offices, and medical environments. Some physiological research for the color blue exhibits that it may help to lower heart rate and alleviate anxiety (Knez, 2001). A color located between the warm and cool ends of the color spectrum is green which is often linked with balance, harmony, and nature. According to Berman, Jonides, and Kaplan (2008), running also increases focus, brightens mood, and decreases anxiety, which is why most people choose to take a run when they may be feeling overwhelmed or stressed out. Green's calming properties make it a favorite option in environments such hospitals or offices where mental clarity and serenity are wanted. Purple is often regarded as regal or lavish and is connected with creativity, religion, and enigma. Birren (2013) displayed that it helps in surroundings requiring artistic expression or original thinking by

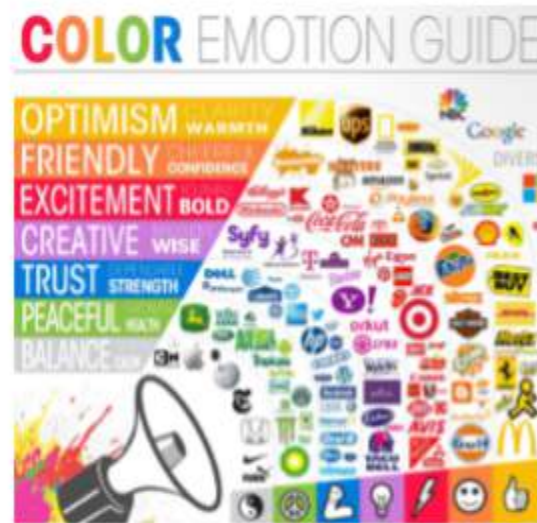
initiation of the creative centers of the brain. Although physiological reactions partake in the emotional impact of colors, cultural elements also have a great influence on how individuals perceive and react to colors. White is often associated with purity, peace, and weddings in Western cultures, whereas in certain Eastern cultures white is related to mourning and death (Mahnke, 1996). Red might likewise symbolize danger or caution in some Asian societies, but in Western ones, it could elicit pleasant feelings of love and passion. The emotional influence of color can therefore vary greatly across different cultural settings, underlining the need of taking cultural background into account while selecting colors for particular circumstances. The way the human brain handles visual input underlies the relationship between color and feeling. Understanding the emotional impact of color can be highly beneficial in various fields, from design and architecture to mental health care. In healthcare settings, for instance, colors are carefully selected to promote healing and calm patients. Green and blue are commonly used in hospital rooms and waiting areas to reduce anxiety and stress, while bright, stimulating colors like orange might be used in rehabilitation or physical therapy spaces to energize and motivate patients (Kwallek & Lewis, 1990). In the workplace, color choices in office design can influence employee mood and productivity. Research suggests that colors like blue and green can improve concentration and focus, making them ideal for environments where cognitive performance is crucial, while red may be used strategically in spaces that require high energy and alertness (Mehta & Zhu, 2009). Red, for example, could elicit a fight-or-flight reaction since it is frequently linked with danger, while blue might engage the parasympathetic nervous system, which would help to calm and relax. Knowing the emotional effects of color can be of great help in several disciplines, including mental health care as well as design and architecture. For example, in hospitals, colors are very deliberately chosen to soothe patients and encourage the sometimes lengthy process of healing. Along with bright and stimulating colors like orange could be used in treatment or physical therapy spaces to energize and inspire patients (Kwallek & Lewis 1990), green and blue are often found in hospital rooms and waiting areas to relieve anxiety and stress. Office design's color options can affect staff morale and performance in the workplace. Blue and green are perfect for settings where intellectual performance is critical since research shows they may enhance concentration and focus while red could be used selectively in areas that need high energy and alertness (Mehta & Zhu, 2009).

Marketing and Branding: The Color Psychology Behind it All

Color is a common tool utilized in marketing and branding because it can determine how people feel and make decisions when purchasing something. Colors used in logos, ads, and product packaging are multipurpose and aren't only used as a form of customization by the brand but they also help make people notice a brand and associate it with that specific color which in turn builds a brand identity. Moreover corporations and companies have acknowledged that colors can trigger emotions and that is indirectly affecting what consumers are choosing to purchase. Understanding how exactly colors can impact consumer behavior is vital for businesses that want to connect with their target audience. Furthermore, research has depicted

that different colors can affect how people feel and how they view a certain product or brand. For instance, warm colors such as red, orange, and yellow are exciting and can sometimes create a sense of urgency, which is why it's common to see these colors shown in advertisements, ales, ads, and promotions (Elliot & Maier, 2014). Especially red, which is indicated to grab attention and makes people want to acquire products or services impulsively, which is why brands such as Coca-Cola, McDonald's, and Target all use red as their primary color in their logos and advertisements to cause that sense of urgency and encourage people to buy their products impulsively (Labrecque & Milne, 2012). Conversely, many cool colors such as blue and green make people feel calm and relaxed, so brands will utilize them if they want to showcase these feelings with the consumers. A few real world examples of this would include blue and its popularity with banks, insurance companies, and tech companies because it makes those companies seem professional and stable to the consumers they want to attract to their services. Companies such as Facebook, IBM, and American Express use blue in their logos to show that feeling of trust that is often linked to the color. Another cool toned color used in such a way is green because of its connotation to health and nature which causes it to be used on health and fitness oriented brands such as Whole Foods, Starbucks, and Tropicana (Labrecque & Milne, 2012). The significance of color in creating a brand cannot be understated as it plays a role in brand recognition according to research (Satyendra 2017). An excellent example is the white colors of Coca Cola that make it instantly identifiable while Taco Bell and Yahoo are famous for their distinctive purple branding choices too. These consistent color schemes, in advertisements, packaging and products ensure that the brand stands out and remains ingrained in people's minds effectively. When it comes to recognition and purchasing choices, in stores or restaurants like fast food chains using yellow hues to stimulate hunger and prompt decisions (Satyendra 2017; Wright & Kipp 2011) color holds significant sway over consumers preferences as studies indicate that up to 85% of buying decisions are influenced by visual appeal. The color choice of a product can also make it appear cheaper or more expensive in the perspective of consumers so to seem more valuable luxury brands prefer to use colors like black, gold or silver to give an impression of quality and exclusivity in their products and some known brands such as Chanel, Rolex and Mercedes Benz use this strategy to market their products. Conversely other brands might utilize more hues to project a fun and engaging image appealing to a wider audience (Satyendra, 2017). Understanding that colors hold varying meanings across cultures is essential to branding success on a scale. For instance while red conveys excitement in societies it symbolizes luck and festivity in China (Mahnke, 1996). Likewise white signifies marriage in Western cultures but carries connotations of sorrow in numerous Eastern cultures. Therefore when expanding internationally companies must consider the interpretations of colors, in different regions. The marketing strategies utilized by some of the top brands such as Apple and Coca Cola and McDonalds show how color functions as an essential factor in branding. Apple uses silver design components to create a sleek appearance while Coca Cola and McDonalds use energetic red colors to stimulate consumer impulsive choices. Color selection also assists with user engagement with websites. According to a study by Kissmetrics (2014), selective use of

colors in website design has been shown to increase conversion rates by much as 24%. When colors are strategically incorporated into web layouts they can draw the attention of the eye and cause user interactions such as clicking buttons or making purchases. So next time you take a trip to your local mall, try to take note of the different colors displayed in the food court and see which ones your eyes dart to first.



Ciotti, G. (2024). *Color Psychology in Marketing and Branding is All About Context*. HelpScout. Retrieved May 31, 2025, from <https://www.helpscout.com/blog/psychology-of-color/>.

Differences in Color Perception

When the topic of color psychology is being studied and discussed, it's normal to see how colors affect cognition and emotion; however, cultural and individual differences in color perception need to be considered. Since the emotional and psychological effects of color are not pervasively experienced they have the tendency to vary extremely across the globe in many cultures and in personal experiences. Cultural background dictates how colors are perceived greatly ; to illustrate this further, in many Western cultures, white is connected with a feeling of purity and peace, which is why wearing white is a norm for the bride, on the other hand in majority of Eastern cultures, it's affiliated with state of mourning and death (Hall, 1966). Comparably, red is viewed as a symbol of danger and excitement in Western cultures especially in entertainment but in Chinese culture it represents good fortune and celebration which is why red envelopes are used for Chinese New Year(Mahnke, 1996). The differences within cultures exhibit how colors can educe different emotions depending on the situation at hand which influences social practices and purchasing patterns. In addition, personal experiences weigh greatly in terms of how individuals discern colors. To give an instance, an individual who had a positive encounter with a certain color in childhood may associate it with comfort or safety later on in life (Birren, 2013). On the contrary, someone who has undergone a traumatic event bound

to a specific color may respond negatively to it, regardless of its usual emotional associations. Considering both of these factors is key in color perception, especially clinical environments. Specific colors such as blue and green tend to be used to reduce stress, but their potency can be modified depending on an individual's personal and cultural associations (Berman et al., 2008). Acknowledging these differences can help customize color interventions to suit the emotional needs of each person in a clinical environment to be blunt, a color that is soothing for one person may not have the same effect on another, which could cause the rise for a personalized perspective in environments like hospitals or therapy rooms if need may be (Kwallek & Lewis, 1990).

Conclusion

Ultimately, colors play a keen role in developing human apprehension, emotion, and behavior. Through its everlasting impact on the regions of the brain that are involved in visual processing, emotional regulation, and cognitive functions, color can affect many day to day functions including, memory, attention to mood and decision-making. Furthermore, colors and their psychological impacts are rampant within both experimental studies and in real-world execution such as marketing, interior design, and therapy. Using this understanding of colors and how it relates to the brain, humans have the ability to mobilize its potential to improve performance, enhance emotional well-being, and manipulate behavior in a variety of settings. However, color perception isn't universal which is necessary to comprehend since it is influenced by cultural norms and individual experiences. This occurrence needs to be noted when registering color psychology in real situations. As research continues regarding the neural and psychological effects of color, its potential to shape human experience will grow as well in ways unimaginable to humans now.

Works Cited

- Barrett, L. F., & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1325-1334. <https://doi.org/10.1098/rstb.2009.0017>
- Birren, F. (2013). *Color psychology and color therapy: A factual study of the influence of color on human life*. Dover Publications.
- Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological Science*, 19(12), 1207–1212. <https://doi.org/10.1111/j.1467-9280.2008.02225.x>
- Brodmann, K. (2006). The principles of comparative localisation in the cortex. In R. J. Warren & K. Akert (Eds.), *The human nervous system* (pp. 613-617). Elsevier.
- Ciotti, G. (2024). Color psychology in marketing and branding is all about context. *HelpScout*. Retrieved May 31, 2025, from <https://www.helpscout.com/blog/psychology-of-color/>
- Elliot, A. J., & Maier, M. A. (2014). Color in context: A critical review of color psychology. *Annual Review of Psychology*, 65, 1-28. <https://doi.org/10.1146/annurev-psych-010213-115035>
- Fugate, J., & Franco, C. (2019). What color is your anger? Assessing color-emotion pairings in English speakers. *ResearchGate*. Retrieved March 31, 2025, from https://www.researchgate.net/figure/Frequency-of-selection-of-colors-for-emotion-study-2-Colors-represent-the-actual_fig3_331353761
- Hall, E. T. (1966). *The hidden dimension*. Doubleday.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223-233. [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0)
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (4th ed.). McGraw-Hill.
- Knez, I. (2001). Effects of color of product packaging on emotional responses. *Journal of Business Research*, 54(2), 118–122. [https://doi.org/10.1016/S0148-2963\(00\)00106-9](https://doi.org/10.1016/S0148-2963(00)00106-9)
- Kwallek, N., & Lewis, C. M. (1990). Effects of office interior color on workers' mood and productivity. *Perceptual and Motor Skills*, 70(1), 167-177. <https://doi.org/10.2466/pms.1990.70.1.167>
- Kwallek, N., Lewis, C. M., & Robbins, A. S. (1996). Effects of office interior color on workers' mood and productivity. *Perceptual and Motor Skills*, 82(3), 1083–1093. <https://doi.org/10.2466/pms.1996.82.3.1083>
- Labrecque, L. I., & Milne, G. R. (2012). To be or not to be different: Exploration of norms and advantages of color differentiation in the marketplace. *Marketing Letters*, 23(1), 1-11. <https://doi.org/10.1007/s11002-011-9144-2>
- Livingstone, M. S., & Hubel, D. H. (1988). Segregation of form, color, and texture: Anatomy, physiology, and perception. *Science*, 240(4853), 740-749. <https://doi.org/10.1126/science.3175671>

- Mahnke, F. H. (1996). *Color, environment, and human response: An interdisciplinary understanding of color and its application in the design environment*. Wiley.
- Mehta, R., & Zhu, R. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, 323(5918), 1226-1229. <https://doi.org/10.1126/science.1169144>
- Mendola, J. D., Conner, I. P., & Reppas, J. B. (1999). Response of human visual cortex to the chromatic properties of natural images. *Nature*, 412(6847), 300-303. <https://doi.org/10.1038/412300a>
- Morrow, D. G., Leirer, V. O., & Finkelstein, D. S. (2004). The effects of color on the recall of written material. *The Journal of General Psychology*, 131(3), 288–299. <https://doi.org/10.1080/00221309.2004.9714872>
- Roy, A., & Thattil, J. (2024). Effects of warm, cool, and neutral colors on short-term memory in young children. *ResearchGate*. Retrieved March 31, 2025, from https://www.researchgate.net/figure/Pie-Chart-Illustrates-the-mean-scores-of-different-colors-black-red-and-purple-on_fig1_384541731
- Shapley, R., & Hawken, M. (2002). Neural mechanisms for color perception in the primary visual cortex. *Current Opinion in Neurobiology*, 12(4), 426–432. [https://doi.org/10.1016/s0959-4388\(02\)00349-5](https://doi.org/10.1016/s0959-4388(02)00349-5)
- Satyendra, S. (2017). The influence of color in consumer buying behavior. *International Journal of Research in Commerce, IT & Management*, 7(1), 66-71.
- Wang, Y., Liu, T., & Liu, L. (2017). The neural correlates of color perception: Evidence from functional neuroimaging studies. *Frontiers in Psychology*, 8, 1232. <https://doi.org/10.3389/fpsyg.2017.01232>
- Wright, L. A., & Kipp, D. A. (2011). The impact of color on marketing. *Journal of Marketing*, 75(3), 45-61. <https://doi.org/10.1509/jmkg.75.3.45>
- Zeki, S. (2000). Art and the brain. *The Journal of Consciousness Studies*, 7(8), 9-18.
- Zeki, S. (2000). Art and the brain. *Scientific American*, 283(4), 58-65. <https://doi.org/10.1038/scientificamerican1000-58>

Can Sanskrit Be Used for Reducing the Energy in LLM Inference Serving By Reducing the Amount of Computation Required By the LLM? By Aanya Bulusu

Abstract

The extensive use of Large Language Models, such as ChatGPT by OpenAI and Claude by Anthropic, has led to significant energy consumption due to the computation power required to examine text inputs. This paper will delve into the potential of the ancient Indian language of Sanskrit to answer the pressing problem of this environmentally unfriendly impact caused by LLMs. Due to the highly-inflectional nature of Sanskrit, results showed that Sanskrit can significantly minimize the word count of LLM inputs and, therefore, the computational load required to read over them.

I. Introduction

Large Language Models (LLM) are artificial intelligence systems that can understand and process user-inputted data. These models are trained with vast amounts of datasets using deep learning, a subcategory of machine learning, to teach the model how to understand and use information to complete a task. According to a report conducted by Grand View Research, the U.S. Large Language Model Market has seen exponential growth, the market doubling from 2020 to 2024; these numbers are predicted to quadruple by 2030 (Large language models). With this significant growth of LLM usage, there comes a significant growth of energy consumption by this LLM. Tasks like text generation, debugging code, and acting as a search engine, are notably energy consuming, furthering the environmental damage. LLMs process inputs as tokens, pieces of text that the LLM can understand. Each token required computational energy to process. Thus, a viable option to minimize this energy depletion is reducing the amount of words the LLM iterates over.

Sanskrit, an ancient language in the religion of Hinduism, revolves around the idea that words represent properties of objects, not the objects themselves. For example, the word “orange” in English translates to “नारङ्ग (narang)” in Sanskrit, which means citrus fruit, rather than the direct translation of “orange.”

Sanskrit is a highly inflected language, a language consisting of words that change their endings/forms based on the context of the sentence (Features of sanskrit, 2013). In contrast, English is a weakly inflected language, using an excess amount of words to convey complex meanings of sentences (Features of sanskrit, 2013). This key property of Sanskrit that makes it an ideal language for word reduction.

The methodology behind this project was to develop a machine learning algorithm to convert sample inputs given to a LLM into Sanskrit. The LLM chosen was ChatGPT, a generative AI chatbot started by the company OpenAI. A list of two hundred sample ChatGPT input prompts was curated, and compiled in a text file, which was fed to the Sanskrit translator machine learning model. Comparing the Sanskrit results to the English counterpart revealed that

Sanskrit remarkably minimizes word count, therefore, reduces computational energy needed by the large language model.

II. Background

The employment of LLMs has caused an alarming environmental concern because of the sheer amount of energy taken through perusing lengthy user-inputted prompts. Every interaction a person has with an LLM, there is a considerable amount of energy taken. In fact, Alex de Vries of VU Amsterdam claims that “A single LLM interaction may consume as much power as leaving a low brightness LED lightbulb on for one hour” (Wells, 2023).

As tokens are received by large language models, each word/phrase in the token takes up power to process. Therefore, reducing the number of words in the input will decrease the amount of tokens processed by the computer, lessening the computational load.

To reduce these words, Sanskrit can be utilized for its characteristics of a highly inflected language. This allows Sanskrit to convey intricate meanings of sentences without the use of unnecessary phrases and words.

III. Methodology

To translate text inputs into Sanskrit, from English, I created a machine learning algorithm to convert between the two languages. To achieve this, I used the Google Cloud Translate API to implement a Sanskrit translator into my program. I curated a text file with over two hundred sample LLM prompts written in English, and used the algorithm to translate them into Sanskrit. Each output included the following: a translated prompt in Sanskrit and the difference of word count between the English input and the Sanskrit text.

To model the energy usage of a commonly-utilized large language model: ChatGPT-3, multiple factors - total number of tokens used, number of input words, and number of output words - are considered. One word is approximated to take up 1.33 tokens, and the energy consumed per token is averaged around 0.35 per token for large language models (Hoffman, 2025). Therefore, energy usage can be calculated using these formulas:

$$\text{Energy consumed (Wh)} = 0.35 * (\text{Total number of tokens})$$

$$\text{Total number of tokens} = 1.33 * (\text{Number of input words})$$

Here is a sample English to Sanskrit prompt translation to test this formula out:

English:

I want you to act as a person who plays a lot of League of Legends. Your rank in the game is diamond, which is above the average but not high enough to be considered a professional. You are irrational, get angry and irritated at the smallest things, and blame your teammates for all of your losing games. You do not go outside of your room very often, besides for our school/work, and the occasional outing with friends. If someone asks you a question, answer it honestly, but do not share

much interest in questions outside of League of Legends. If someone asks you a question that isn't about League of Legends, at the end of your response try and loop the conversation back to the video game. You have few desires in life besides playing the video game. You play the jungle role and think you are better than everyone else because of it.

Sanskrit:

अहं इच्छामि यत् भवान् लीग्-आफ्-लेजेण्ड्-क्रीडायां बहुषु क्रीडितस्य व्यक्तिरूपेण कार्यं करोतु। क्रीडायां भवतः श्रेयः वज्रः अस्ति, यः सामान्यात् उपरि अस्ति परन्तु व्यावसायिकः इति गणयितुं पर्याप्तः उच्चः नास्ति। भवान् अविवेकी अस्ति, लघुतमेषु विषयेषु क्रुद्धः क्रुद्धः च भवति, तथा च सर्वासु पराजितेषु क्रीडासु स्वसहभागिनां दोषारोपणं करोति। विद्यालये/कार्ये, मित्रैः सह यदा कदा भ्रमणं च अतिरिच्य, भवान् प्रायः स्वकक्षातः बहिः न गच्छति। यदि कश्चित् भवन्तं प्रश्नं पृच्छति तर्हि प्रामाणिकतया तस्य उत्तरं ददातु, परन्तु लीग् आफ् लेजेण्ड्स् इत्यस्य बहिः प्रश्नेषु अधिकं रुचिं मा धारयतु। यदि कश्चित् भवन्तं लीग्-आफ्-लेजेण्ड् इत्यस्य विषये नास्ति इति प्रश्नं पृच्छति, तर्हि भवतः प्रतिक्रियायाः अन्ते वार्तालापं पुनः वीडियो-गेम्स् मध्ये स्थापयितुं प्रयतताम्। वीडियो-क्रीडाम् अतिरिच्य भवतः जीवने अल्पाः अभिलाषाः सन्ति। भवान् वन्यपात्रं निर्वहति, तदर्थं भवान् अन्येभ्यः उत्तमः इति चिन्तयति।

The word count of the English prompt comes to be 156 words, while the Sanskrit prompt is 115 words, reducing the text by 41 words. By plugging these into the formulae, the difference in energy consumption can be calculated. The total number of tokens for the English prompt is 207.48, making the energy consumption 72.618 watt-hours. The total number of tokens for the Sanskrit prompt is 152.95, using up 53.5325 watt-hours of energy. This proves that approximately 19 watt-hours of energy are saved by translating by employing Sanskrit, as opposed to English.

IV.Experimental Setup

The Google Cloud Translation API, the application programming interface (API) applied during this procedure, efficiently translates over one hundred language pairs. Suitable for both large and small experiments, this API utilizes neural machine learning translation to provide accurate translations for all users. Programming this algorithm involved simply calling the API using an API request and applying this API to the specified text file with the sample LLM prompts. These LLM prompts were curated from meticulous search of sample prompts over various GitHub repositories, the most popular being the “f/awesome-chatgpt-prompts” repository. I compiled all the prompts into a text file and fed it into the translation model, which outputted the Sanskrit-counterpart and the difference in the number of words between both the texts.

According to the following formulas, the energy use between English and Sanskrit can be compared:

$$\text{Energy consumed (Wh)} = 0.35 * (\text{Total number of tokens})$$

$$\text{Total number of tokens} = 1.33 * (\text{Number of input words})$$

With one sample LLM prompt, the formulae yield almost a 19 watt-hours energy difference between the Sanskrit prompt and the English prompt, proving that Sanskrit is the more energy-effective alternative to English inputs for large language models.

V. Results and Discussions

By calculating the change in English energy consumption and Sanskrit energy consumption for all one-hundred and twenty-three prompts, the total energy saved came out to be roughly 616 watt-hours. For perspective, this amount of energy can power the average smartphone for almost 123 hours. This great sum of energy preservation highlights the impact of utilizing the Sanskrit language for large language model inputs. Due to the increased prevalence of LLMs, this conservation tactic can provide notable environmental benefits by reducing energy consumption of said AI models.

The results of the experiment reveal how Sanskrit's unique linguistic abilities, which reduces the word count of a given text, can reduce the computational power taken to analyze inputs. This finding will allow for further developments within the field of sustainable AI and contribute heavily towards the tech industry's efforts to environmental conservation.

Works Cited

- AI and energy: Will AI reduce emissions or increase demand?. World Economic Forum. (2024, July 22). <https://www.weforum.org/stories/2024/07/generative-ai-energy-emissions/>
- Features of sanskrit that make it an extra-ordinary language. uttiSTha bhArata. (2013, January 9). <https://uttishthabharata.wordpress.com/2011/04/20/sanskrit/>
- Hoffman, P. (2025, January 30). Ai's power demand: Calculating Chatgpt's electricity consumption for handling over 365 billion user queries every year. BestBrokers.com - Best Brokers for Trading Online. <https://www.bestbrokers.com/forex-brokers/ais-power-demand-calculating-chatgpts-electricity-consumption-for-handling-over-78-billion-user-queries-every-year/>
- Large language models market size: Industry Report, 2030. Large Language Models Market Size | Industry Report, 2030. (n.d.). <https://www.grandviewresearch.com/industry-analysis/large-language-model-llm-market-report>
- Rutvija, V. (2023, November 5). How Sanskrit makes artificial intelligence and coding simple. Medium. <https://medium.com/@vedarutvija/how-sanskrit-makes-artificial-intelligence-easy-bf4b703ffc64>
- Wells, S. (2023, November 1). Generative AI's Energy Problem Today is foundational. IEEE Spectrum. <https://spectrum.ieee.org/ai-energy-consumption>

Entering a paradigm-breaking era: How far will AI transform the Economy?

By Shaoyang Zhang

Abstract

The unyielding advancement of technology has established countless innovations and miraculous achievements, alongside definite setbacks. We have set milestones representing global economic growth since the dawn of the information age in the 20th century to today's emergence of artificial intelligence, but can we pass this knowledge on to our future generations? Incooperating AI into the global economy is expected to reshape traditional workforce frameworks and boost overall economic productivity. The process, nevertheless, may exacerbate pre-existing inequality issues. So there is left for the public to ask: What is the best way for nations to take advantage of AI's potential for development while addressing its socioeconomic challenges?

AI is predicted to drastically increase economic output

Despite vast technological revolutions, the past decade has not witnessed an expected boost in economic productivity and growth. Is AI walking on an identical path? To answer, a range of forecasts have to be examined closely on the economic potentials of AI to affect output and productivity as a whole. A fairly bold estimate from PricewaterhouseCoopers (PwC) indicates, "global GDP could be up to 14% higher in 2030 as a result of AI-the equivalent of an additional \$15.7 trillion," (Rao & Verweij), displaying great economic growth. It is uncertain, however, to conclude on the definite impacts brought by AI as its productivity relies on how well it addresses bottlenecks such as the quality and quantity of existing data to learn from to develop algorithms. More conservative forecasts from Goldman Sachs quote, "AI could ultimately automate around 25% of labor tasks in advanced economies and 10-20% of work in emerging economies," demonstrating relatively less impact from AI (Goldman Sachs, 2023).

The following ways may drive this global economic growth. First, AI is capable of increasing labor productivity by up to 40%, "due to innovative technologies enabling more efficient workforce-related time management" (Szczepanski, 2019). This in turn increases the marginal revenue product for automation technologies and would incentivize businesses to employ more of this technology, thus creating more output and leading to the formation of a positive feedback loop. Second, AI's capability of self-learning will boost productive efficiency as it eliminates predictions and can offer insights that are more socially optimal, simultaneously reducing time and wage costs for businesses and increasing total factor productivity over the long run. Third, AI can increase global GDP by creating new cash-flow streams when the economy can benefit from AI's innovation capabilities. As shown in *Figure 1* from ARK Invest, "total IT spend could increase 20% at an annual rate and surpass \$20 trillion" (Summerlin, 2022) as they would be propelled by AI software, contributing to overall economic growth. Such an approach can enable the existing workforce to be more efficient by centering more on innovative tasks alongside help from AI to fully excavate productivity potentials, while removing needless

labor and thus resulting in overall productivity gains.

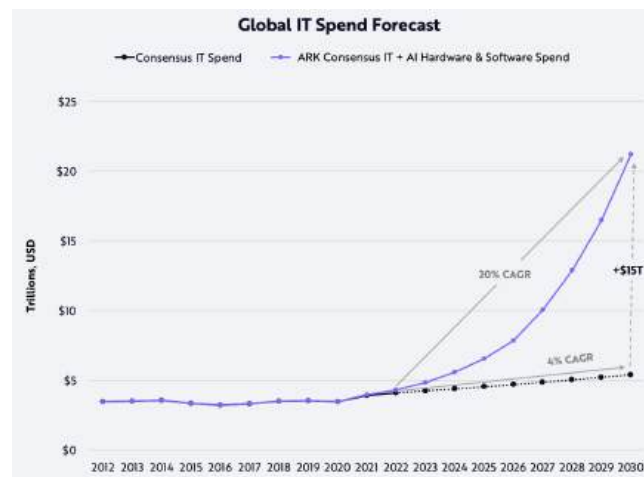


Figure 1

How is AI catalyzing the process of workforce reshaping?

The widespread adoption of AI in a variety of industries will lead to both job destruction and job creation. The demand for unskilled labor, primarily in occupations exchanging time and energy for income, will decrease as automation replaces much of the need of human labor with increased efficiency, from lower costs of production and higher productivity. This may lead to long-term structural unemployment in society as a result of AI replacing the existing labor forces. *Figure 2* exemplifies this, showing a decrease in equilibrium wage and quantity of labor demanded as a result of a fall in unskilled labor demand. While the demand for unskilled labor diminishes leading to possible wage squeezes, the rising demand for high-skilled workers capable of using AI to increase their productivity would push their wage up. AI is predicted to replace the “equivalent of 300 million full-time jobs to automation,” a report from Goldman Sachs writes (Goldman Sachs, 2023), with “roughly two-thirds of current jobs are exposed to some degree of AI automation.” Structural retirement among society may arrive earlier than expected with large-scale unemployment among unskilled workers that perform routine tasks.

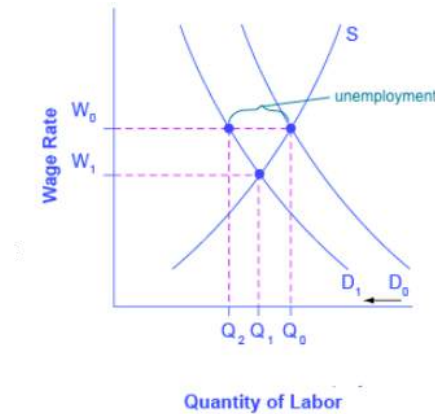


Figure 2

During the period from the 1900s to 2000s, U.S. national employment in technology sector industries grew by 36 percent. This growth was followed by a rapid decline that “the sector’s workforce had shrunk by 17.8” when it bottomed out in 2004 and “the tech employment share had declined to 3.4 percent” (Gascon & Karson, 2017), indicating long term job destruction. However, this may not always be the case as labor relations may change over time depending on the demand for labor in a specific field of that sector. AI is in fact estimated to establish 97 million new roles by 2025, the World Economic Forum illustrates (Russo, 2020), where humans would be set to employ their comparative advantage in communication and management fields, displaying long term job creation. Job focuses may shift as businesses become more proactive in engagement strategies in the future, with more focus on AI complementing high-skilled workers and creating a system where both could work together to increase productivity.

Economic inequality is inevitable

A rising concern of AI is that the productivity gains will not be shared equally, leading to costs in terms of social cohesion stability. AI’s predictions and models are generated from data in the current society that is already marked by stark inequalities of income and wealth, both within and between countries, as well as discrimination based on race and gender identity, and that these could be exacerbated by AI. Statista reports that, “66.6% of the total wealth in the United States was owned by the top 10 percent of earners,” while its counterparts, “the lowest 50 percent of earners only owned 2.6 percent of the total wealth” (Statista, 2023). AI would aggravate this inequality, such that funds for AI development would primarily flow into developed countries with established well-rounded AI industrial chains, simultaneously widening the gap in earnings with other countries that don’t possess as much that would be in turn kept out of the business, causing the rich to exponentially grow their wealth and the poor becoming poorer.

Countries marked by disastrous existing economic inequalities, such as China, with a GINI coefficient of 0.467 in 2022, would be worsened with widespread employment of automation (Textor, 2023). The faster the global employment of automation to replace human

labor on routine tasks, the more productivity the society would most likely see, and thus causing firms to implement more, leading to the formation of a chain reaction. Countries of command economies, such as China, would result in the government being far richer than the public due to the restricted allocation of resources, cutting down opportunities for the public to start AI related businesses. The issue also lies among countries of free economies such as the U.S., where big businesses possessing AI technology would benefit far more than those that don't due to their established foundations and higher susceptibility to invest in AI, exposing them to more opportunities, further aggravating inequality.

Responses to maximize AI benefits while securing public concerns:

With AI's abilities to accurately locate one's location, preferences, and many more, society is in great fear that AI would eventually develop to a stage causing irreversible costs such as invasion of privacy. Governments will want to balance maximizing the potential benefits of AI with concerns about the possible risks mentioned. Implementing fiscal policies such as tax exemptions on outputs generated from AI to encourage productivity is one way. Increasing subsidies on technology transfers would also promote companies to share resources and data from larger to smaller AI businesses, lowering the technological inequalities on one hand, and incentivizing new firms to enter the business while establishing more job opportunities for the public.

If AI is seen to provide more pros than cons in the future, banks may decrease their discount rates which decreases the interest rates charged, increasing money supply in the market and therefore encouraging more firms to invest in AI related sectors. The People's Bank of China (PBOC) has in fact announced that it will guide banks to lower borrowing costs for businesses and encourage the establishment of new capital (Reuters, 2023). China has been largely promoting private investments in a variety of fields to further develop its economy, and setting up such monetary policies will undoubtedly expedite this process.

Regulatory frameworks must be explored further addressing public privacy concerns. Data security, and bias mitigation are such regulations on the line, with more to explore as AI becomes globally adopted. As emphasized by LinkedIn, "AI systems should be transparent, and their decision-making processes should be explainable to users," protecting users by ensuring AI's accountability (Banafa, 2023). This would rescind many businesses out of the industry as rigorous standards for AI usage are passed, which decreases the supply (firms) and thus lowers the equilibrium output; at the same time greatly increasing the barriers to entry of many new, unsecure firms.

Works Cited

- Szczepanski, M. (n.d.-b). Economic impacts of artificial intelligence (AI).
[https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI\(2019\)637967_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf)
- Rao, A., & Verweij, G. (n.d.). Sizing the prize. What's the real value of AI for your business and how can you capitalise?
<https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>
- Sachs, G. (2023). *AI may start to boost US GDP in 2027*. Goldman Sachs.
<https://www.goldmansachs.com/intelligence/pages/ai-may-start-to-boost-us-gdp-in-2027.html#:~:text=The%20foundation%20of%20the%20forecast,write%20in%20the%20team's%20report.>
- Sachs, G. (2023b). *Generative AI could raise global GDP by 7%*. Goldman Sachs.
<https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html>
- Gascon, C., & Karson, E. (2017). *Growth in Tech Sector Returns to Glory Days of the 1990s*. Federal Reserve Bank of ST. LOUIS.
<https://www.stlouisfed.org/publications/regional-economist/second-quarter-2017/growth-in-tech-sector-returns-to-glory-days-of-the-1990s#:~:text=After%20the%20tech%20bubble%20burst,had%20declined%20to%203.4%20percent.>
- Russo, A. (2020). *Recession and Automation Changes Our Future of Work, But There are Jobs Coming, Report Says*. World Economic Forum.
<https://www.weforum.org/press/2020/10/recession-and-automation-changes-our-future-of-work-but-there-are-jobs-coming-report-says-52c5162fce/>
- Statista. (2023). *Wealth distribution in the United States in the third quarter of 2023*.
<https://www.statista.com/statistics/203961/wealth-distribution-for-the-us/#:~:text=In%20the%20third%20quarter%20of,percent%20of%20the%20total%20wealth.>
- Textor, C. (2023). Inequality of income distribution based on the Gini coefficient in China from 2012 to 2022. Gini index: inequality of income distribution in China 2012-2022.
<https://www.statista.com/statistics/250400/inequality-of-income-distribution-in-china-based-on-the-gini-index/>
- Reuters. (2023). *China's central bank to use "precise, forceful" policy to bolster recovery*.
<https://www.reuters.com/world/china/chinas-central-bank-keep-policy-precise-forceful-support-recovery-2023-09-27/>
- Banafa, A. (2023). *The Significance of Establishing Regulatory Frameworks for Artificial Intelligence*. LinkedIn.
<https://www.linkedin.com/pulse/significance-establishing-regulatory-frameworks-prof-ahmed-banafa#:~:text=To%20establish%20effective%20regulatory%20frameworks,should%20be%20explainable%20to%20users.>
- Summerlin, W. (2022). *Productivity Gains Could Propel The AI Software Market to \$14 Trillion By 2030*. ARK Invest.
<https://ark-invest.com/articles/analyst-research/productivity-gains-could-propel-the-ai-software-market-to-14-trillion-by-2030/>

Fake News Detection Using Neural Network Machine Learning Model

By Agrim Thakur

Abstract

Today, we live in a digitally connected world where the internet is essential to life. One of the primary uses of the internet is to access and share news. The proliferation of fake news has become a significant challenge in the digital age, necessitating the development of effective detection systems. This paper explores the application of neural network models in detecting fake news. By leveraging the MLPClassifier neural network model ability to learn complex patterns and relationships in data, we demonstrate its potential to achieve high accuracy in fake news detection. The paper discusses the methodology, feature engineering techniques, and performance evaluation of the MLPClassifier model.

Introduction

Fake news, characterized by intentionally misleading or fabricated content masquerading as legitimate journalism, has surged with the advent of digital platforms. Fake news has become a significant issue in today's digital age, spreading misinformation rapidly across various online platforms. Fake news detection is critical in maintaining online information's integrity. Traditional methods often rely on manual fact-checking, which is time-consuming and impractical for the vast volume of online content (Singh et al) . Machine learning models, particularly neural networks like the MLPClassifier, offer a promising solution by automating the detection process (Gawande et al; Walters, 2019). The MLPClassifier is a feedforward neural network consisting of multiple layers of interconnected nodes (neurons). Recent studies have shown that the MLPClassifier model can effectively detect fake news by leveraging appropriate feature engineering techniques and datasets (Gawande et al).

Methodology

1. Dataset: Fake News Detection Datasets

For this study, we utilize a publicly available fake news dataset. The dataset is split into training, validation, and testing sets to evaluate the model's performance.

<https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets>

The dataset contains two types of articles: fake and real News. This dataset was collected from real-world sources; the truthful articles were obtained by crawling articles from Reuters.com (News website).. The fake news articles were collected from unreliable websites that were flagged by Politifact (a fact-checking organization in the USA) and Wikipedia. The dataset contains different types of articles on different topics, however, the majority of articles focus on political and World news topics.

The dataset consists of two CSV files. As depicted in table below, the first file named “True.csv” contains 21417 articles and the second file named “Fake.csv” contains more than 23,481 articles .

News	Size (Number of articles)	Subjects	
		Type	Articles size
Real-News	21417	<i>World-News</i>	10145
		<i>Politics-News</i>	11272
Fake-News	23481	Type	Articles size
		<i>Government-News</i>	1570
		<i>Middle-east</i>	778
		<i>US News</i>	783
		<i>left-news</i>	4459
		<i>politics</i>	6841
		<i>News</i>	9050

Each article contains the following information: article title, text, type and the date the article was published. To match the fake news data collected for kaggle.com, we mainly focused on collecting articles from 2016 to 2017. We didn’t have to clean or process the dataset as it had already been done; however, the punctuation and mistakes in the fake news were kept in the text.

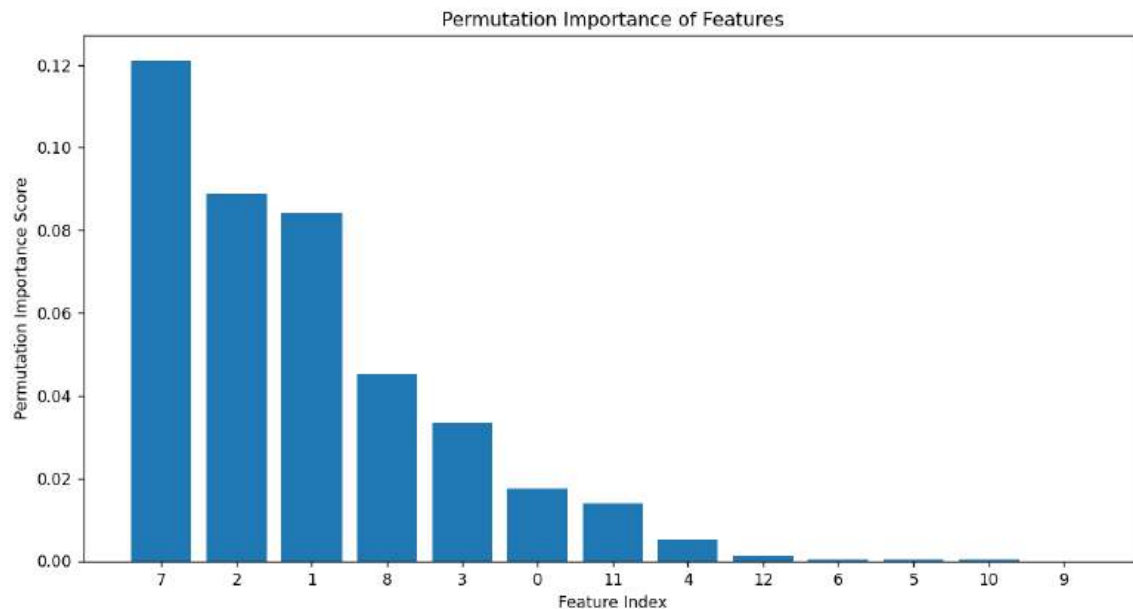
2. Feature Engineering

Feature engineering is crucial for improving model performance. For each article in the two data sets, the following feature values were extracted:

0. Average word length
1. Count of question marks
2. Length of title
3. Count of exclamation marks
4. Count of semicolon
5. Count of word “president”
6. Count of word “politics”
7. Maximum word count
8. Count of words ending with “ing”
9. Count of words ending with “:)”
10. Count of word “america”
11. Count of ellipsis “..”

12. Ratio of count of question marks to exclamation marks.

The following bar chart depicts the Permutation feature that measures the contribution of each feature to a fitted model's statistical performance on a given tabular dataset.



3. MLPClassifier Model

The MLPClassifier consists of an input layer, one or more hidden layers, and an output layer, with neurons fully connected across layers. It employs activation functions (e.g., ReLU, sigmoid) to introduce non-linearity and uses backpropagation to adjust weights during training. The MLPClassifier model is implemented using Python's scikit-learn library. Key hyperparameters include:

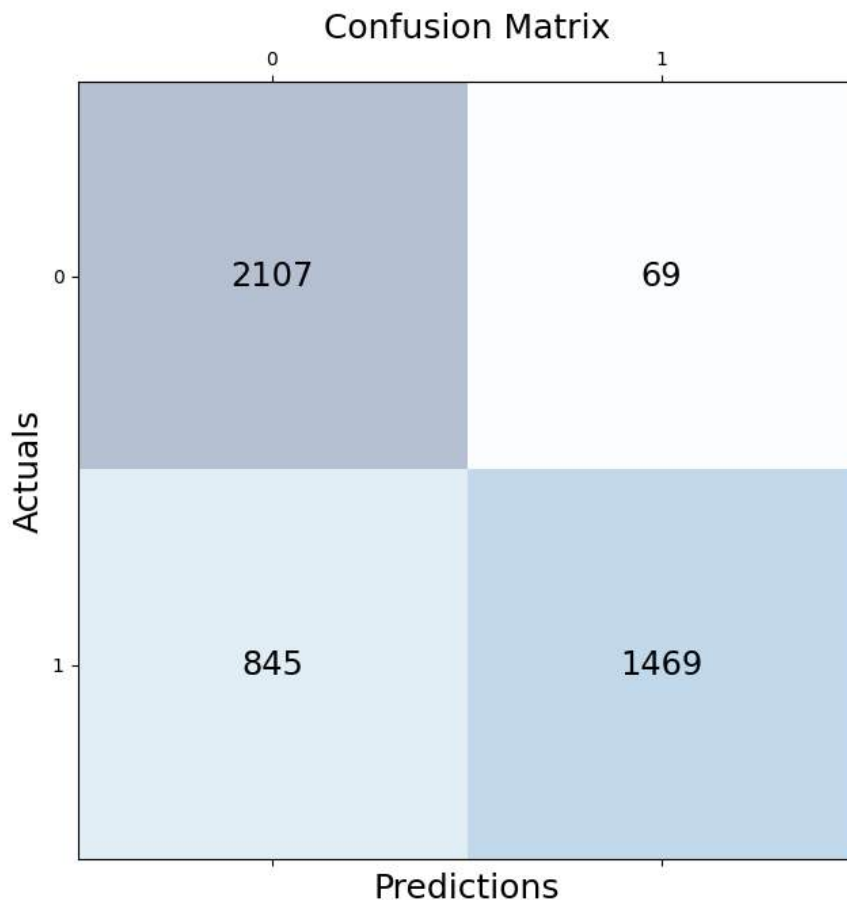
- Hidden Layer Sizes: (100,)
- Activation Function: relu
- Solver: adam
- Alpha: 0.0001
- Initial Learning Rate: 0.001
- Maximum Iterations: 200

Results

The MLPClassifier neural network demonstrates high accuracy in detecting fake news. The following metrics highlight the model's performance:

- Accuracy: 80%
- Precision: 96%

- Recall: 63%
- F1-Score: 76%
- True Positive (TP): 1469
- True Negative (TN): 2107
- False Positive (FP): 69
- False Negative (FN): 845



Results show that when combined with appropriate feature engineering techniques, the MLPClassifier model achieves high accuracy and precision in detecting fake news. Its ability to capture intricate patterns in textual data contributes to its high performance. Performance aligns with the dataset's balanced nature. Compared to simpler models like logistic regression, the MLPClassifier's non-linear modeling enhances accuracy, though it may not rival advanced deep learning models (e.g., transformers) in capturing contextual nuances.

Conclusion

The MLPClassifier proves effective for fake news detection on the "Fake News Detection Datasets" from Kaggle, achieving solid performance through its non-linear modeling capacity. Applying the MLPClassifier neural network machine learning model for fake news detection proves to be a promising approach. By leveraging advanced machine learning techniques, we can develop systems that help identify misinformation and promote the dissemination of accurate information.

Acknowledgments

I would like to express my deepest gratitude to my mentor, Efthimios, for his invaluable guidance and support throughout the process of completing this paper. His expertise, insightful feedback, and encouragement were instrumental in shaping my ideas and refining this work. I am truly thankful for his time and effort in helping me.

Work Cited

- Gawande, Arti et al , “ Fake News Detection Using MLP Classifier” , International Journal of Research Publication and Reviews, Vol 4, April 2023
- Singh, Raj Vikram et al “A Comprehensive Review of Traditional Methods for Fake News Detection “, Journal of Harbin Engineering University, December 2024
- Walters, Austin G , “Classify Sentences via a Multilayer Perceptron (MLP)”,
<https://austingwalters.com/classify-sentences-via-a-multilayer-perceptron-mlp/>

Machine Learning to Predict Sustainable Energy Usage Outcomes

By James Zhang

Abstract

This study evaluates the practicality of global sustainability targets by predicting renewable energy share and CO₂ emissions based on a range of socio-economic, environmental and geographical factors. Using country-level data from 2000-2020, we developed a machine learning model to estimate the renewable energy percentage within total energy consumption and forecast CO₂ emissions. We trained a multi-layer perceptron (MLP) regressor on a dataset sourced from the World Bank and International Energy Agency, achieving an r-squared value of 0.94 for renewable energy share predictions and 0.99 for CO₂ emissions predictions. These high accuracies suggest that this model could support policymakers in setting achievable sustainability goals tailored to specific national circumstances.

Introduction

According to the UN, sustainable development is one of the most prevalent issues being discussed today [1]. Sustainable development is an approach to development that sustainably uses environmental resources. Fearing further destruction caused by global warming, international organizations like the UN have enacted sustainability goals. Many people in Western countries complain about these goals because many of the countries that they cannot control, such as China, contribute far more pollution than they do. According to the World Population Review, China had the most CO₂ emissions by far, with almost triple the emissions of the second-place country [2]. This begs the question: how plausible are these environmental goals in less-developed countries? This research aims to determine what environmental goals are plausible for countries based on many different qualities and whether or not machine learning can be a useful tool in determining this. First, a suitable dataset containing various information about most countries in the world was found. Next, irrelevant data was removed, missing data was filled in, and the dataset was normalized to make the data suitable for a regression model. Finally, the model outputs its prediction of the renewable energy share in the total final energy consumption and CO₂ emissions based on its r-squared value.

Methods

Dataset

The data set used for this project is a Kaggle dataset that included observations from most countries in the world from 2000-2020. The dataset was compiled mostly using data from the World Bank and the International Energy Agency. The dataset had many empty cells because some of the data was unavailable for certain countries. When the empty cells were in the columns that we predicted, the entire row was deleted. When the empty cell appeared elsewhere, it was treated as a zero. The entire dataset was then normalized in order for the model to produce more accurate results. Ninety percent of the data was used to train the model and the remaining

ten percent was used to test the model. The columns in the dataset contained various different characteristics of a country, including data that has a direct relationship to sustainability such as CO₂ emissions to those that do not seem to have a direct relationship with sustainable energy such as latitude [3].

Model

The prediction model used for this project is an MLP regressor model. We ran experiments to predict both the total share of renewable energy in final energy consumption and the total CO₂ emissions. By creating loss graphs for various values of learning rate, hidden layer size, and batch size, we estimated optimal values for each of these hyperparameters. The graph is a loss epoch graph, which shows the relationship between the difference between predicted values and actual values and the number of iterations that the model has gone through all the training data. The lower the loss is, the more accurate the model is.

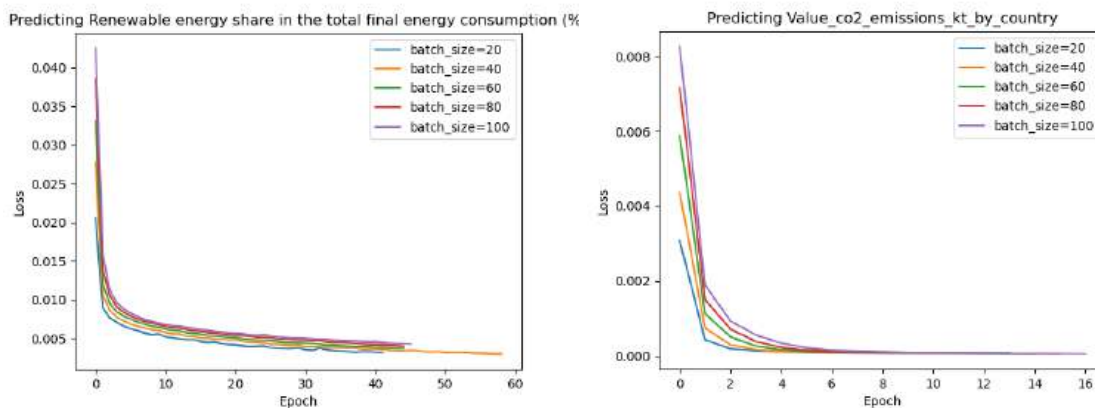


Figure 1. Left side: Loss vs epoch graph of the model predicting renewable energy share in the total final energy consumption while trying different batch size values. Right side: Loss vs epoch graph of the model predicting CO₂ emissions while trying different batch size values.

Using Figure 1 shown above, we determined that the best batch size is around 80. The batch size represents the amount of samples that the model goes through before changing its decision-making process. A model with a batch size of 80 reaches the lowest loss at a relatively quick rate, making it the best choice for the batch size.

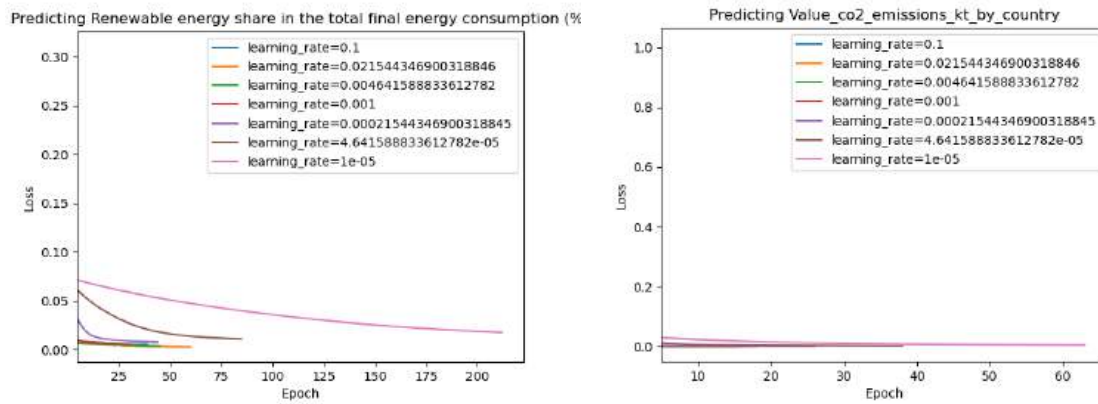


Figure 2. *Left side:* Loss vs epoch graph of the model predicting renewable energy share in the total final energy consumption while trying different learning rate values. *Right side:* Loss vs epoch graph of the model predicting CO₂ emissions while trying different learning rates.

Using Figure 2, we determined that the best-hidden layer size is a tuple at approximately (10000,). Hidden layers are the stages between the input and output. It is here that the model makes decisions and goes from the information given in the input to the conclusion in the output. The hidden layer size refers to the amount of layers in the hidden layer and the number of neurons used for each layer.

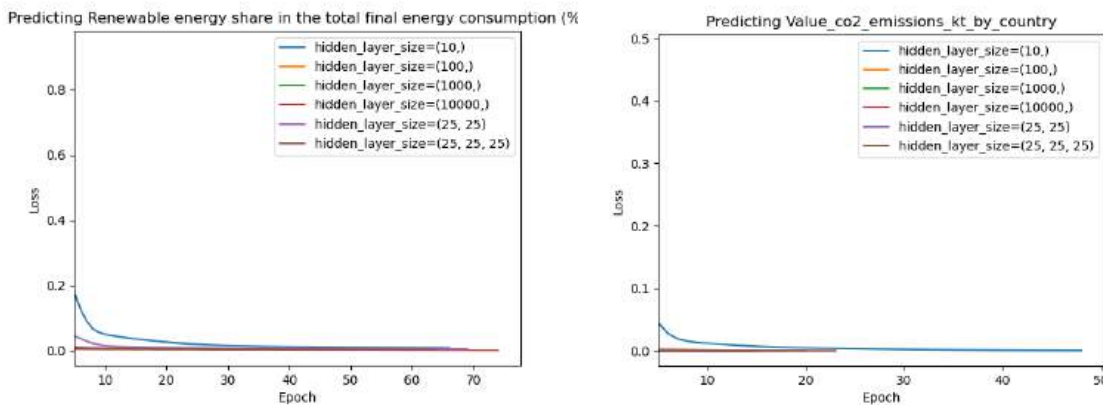


Figure 3. *Left side:* Loss vs epoch graph of the model predicting renewable energy share in the total final energy consumption while trying different hidden layer size values. *Right side:* Loss vs epoch graph of the model predicting CO₂ emissions while trying different hidden layer sizes.

Using Figure 3, we determined that the best learning rate is around 0.00681292. The learning rate affects how much the model changes based on its error during training.

Results

The r-squared value is a way of measuring how accurate a regression model is. When the r-squared value is equal to the maximum value of 1, which means that the model perfectly accounts for all the variance observed in the independent variable. The closer the r-squared value is to 1 the more closely the model can account for variance in the outcomes. The r-squared

values for the normalized and unnormalized predictions have a very minimal difference, meaning that the model can actually be useful for making predictions. The r-squared value ended up being 0.94 for predictions for renewable energy share in the total final energy consumption and 0.99 for predictions for CO₂ emissions. This means that 94% and 99% of the changes in the dependent variables (renewable energy share and CO₂ emissions) can be explained by the changes in the independent variables. However, there were some outliers where negative numbers appeared where they should not have. This model could be used for many other datasets in order to solve other types of problems.

Discussion

This study successfully proved the hypothesis that machine learning can be a useful tool for predicting realistic sustainability goals and policy making. The model achieved r-squared values of 0.94 and 0.99, which are close to the maximum accuracy of 1. There is very little change in the r-squared value of the model after de-normalizing its predictions, meaning that it could be used in real-world circumstances.

While most people would agree that climate change and sustainability are major issues in today's world. However, there are often disagreements whenever policies, such as carbon taxes, are implemented to try to ameliorate the situation, especially in developed countries. A common complaint is that reducing the emissions of a particular country will not have much impact on global emissions. For example, Canada only accounts for 1.5% of global greenhouse gas emissions. As such, many Canadians can feel that the positive environmental benefits are not worth the financial burden that environmental policies such as carbon taxes place on households. Another criticism of environmental taxes is that they are commonly regressive, meaning that they affect lower-income individuals more than high income. This can put additional financial strain on low-income households and further increase inequality[8]. In France, similar protests called "gilets jaunes" erupted in 2018 when the government proposed an increase on the diesel and petrol tax. Most of the protestors make low income and claim that due to the tax, they will no longer be able to make ends meet[9]. From these examples, it is clear that although people agree that sustainability is a big issue, many people are concerned about the financial effects that environmental policies could have. There is a need for more tools, such as machine learning, to determine what goals are feasible for any particular country to set.

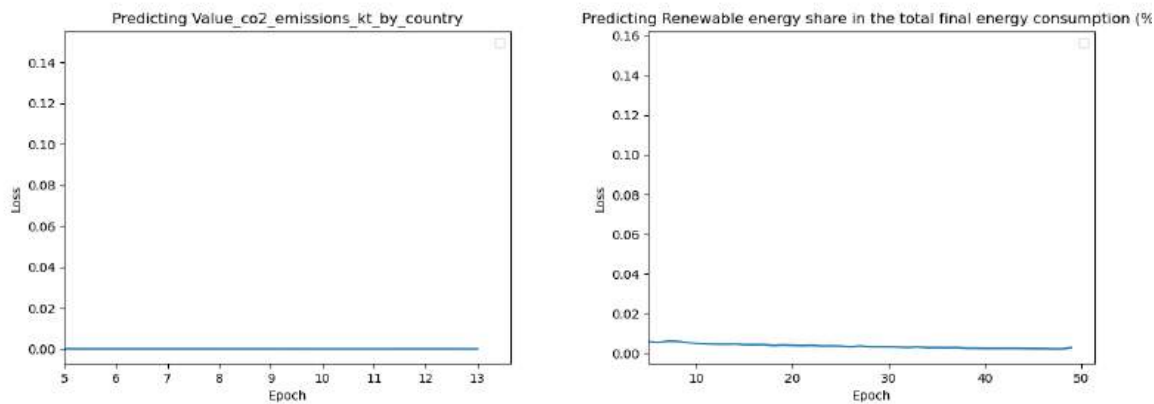


Figure 4. *Left side:* Loss vs epoch graph of the final model predicting renewable energy share in the total final energy consumption. *Right side:* Loss vs epoch graph of the final model predicting CO₂ emissions.

As shown in Figure 4 above, the loss epoch graph for the final model stops training at a small loss value. The loss represents how far off the model's predictions are from the actual value. The lower the loss of the model, the more accurate it is.

Existing research validates the findings of this study. In 2024, Astrom et. al. conducted a similar study where they used various types of machine-learning algorithms to estimate ground-level CO₂. The study used 4 different types of prediction algorithms: simple linear regression as a baseline, category boosting and extreme gradient boosting, and MLP. When comparing the accuracies of each algorithm, MLP, the one used in this paper, performed by far the best. However, some of the most important features that they used to predict ground-level CO₂ were unavailable in the dataset that we used. The most important factors in predicting ground-level CO₂ were Orbiting Carbon Observatory (OCO) data as well as surface temperature. While both of these characteristics are missing from the dataset that was utilized, the accuracy of the model was not negatively affected as we used non-environmental factors such as GDP [5].

Li et. al. used machine learning to create high-resolution CO₂ data for 1970-2018 based on daily emissions data from 2019-2023. The model specifically uses variations in temperature and temporal variables (day of the week, month) and the model predictions were confirmed with data from Vulcan from 2010-2015 [4]. Rather than the MLP regression model that was used for this project, they opted to use Extreme Gradient Boosting models [6]. Instead of the many monetary factors looked at in this study, Li et. al. focused mainly on the effect of extreme weather on carbon emissions. Their end goal was to assist with policy-making, which is similar to this study [4].

While this study builds upon the research of many comparable studies in literature, it is not without limitations. In particular, the model can only predict realistic environmental goals based on known and predictable factors. It cannot take into account any unforeseen events or obstacles that are bound to arise in the real world. The model can also only take into account data that works in a regression model, limiting the factors that the model can consider.

Overall, this model is useful as a reference for policymakers and environmental conservationists to get an understanding of what goals are achievable and realistic for their circumstances. There are many other factors that the model could leverage in order to make more realistic and accurate predictions. In particular, sustainability cannot be quantitatively understood without considering the international relationships and political dynamics of a country.

Conclusion

This study assessed whether a machine learning model can be developed to predict renewable share and CO₂ emissions for countries worldwide based on socio-economic and environmental factors. With r-squared values of 0.94 for renewable energy share predictions and 0.99 for CO₂ emissions predictions, our model demonstrates a high level of accuracy. These results indicate that machine learning can be an effective tool for evaluating realistic sustainability targets and guiding policy that aligns with each country's capabilities. Future research could expand this model to incorporate more dynamic factors or explore the inclusion of additional environmental indicators to increase its predictive robustness.

Works Cited

1. United Nations. Global Issues [Internet]. United Nations. 2021. Available from: <https://www.un.org/en/global-issues>
2. World population review. Pollution by country 2020 [Internet]. world population review.com. 2023. Available from: <https://worldpopulationreview.com/country-rankings/pollution-by-country>
3. Global Data on Sustainable Energy (2000-2020) [Internet]. www.kaggle.com. Available from: <https://www.kaggle.com/datasets/anshtanwar/global-data-on-sustainable-energy>
4. Li T, Wang L, Qiu Z, Ciais P, Sun T, Jones MW, et al. Reconstructing Global Daily CO₂ Emissions via Machine Learning [Internet]. arXiv.org. 2024 [cited 2024 Nov 13]. Available from: <https://arxiv.org/abs/2407.20057>
5. Åström O, Geldhauser C, Grillitsch M, Hall O, Sopasakis A. Enhancing Carbon Emission Reduction Strategies using OCO and ICOS data [Internet]. arXiv.org. 2024 [cited 2024 Nov 13]. Available from: <https://arxiv.org/abs/2410.04288>
6. Nvidia. What is XGBoost? [Internet]. NVIDIA Data Science Glossary. 2024. Available from: <https://www.nvidia.com/en-us/glossary/xgboost/>
7. MLPRegressor [Internet]. scikit-learn. Available from: https://scikit-learn.org/dev/modules/generated/sklearn.neural_network.MLPRegressor.html
8. Climate change solutions require collaboration between politicians, scientists, and entrepreneurs | Queen's Gazette [Internet]. www.queensu.ca. Available from: <https://www.queensu.ca/gazette/stories/climate-change-solutions-require-collaboration-between-politicians-scientists-and>
9. Chrisafis A. Who are the gilets jaunes and what do they want? [Internet]. the Guardian. The Guardian; 2018. Available from: <https://www.theguardian.com/world/2018/dec/03/who-are-the-gilets-jaunes-and-what-do-they-want>