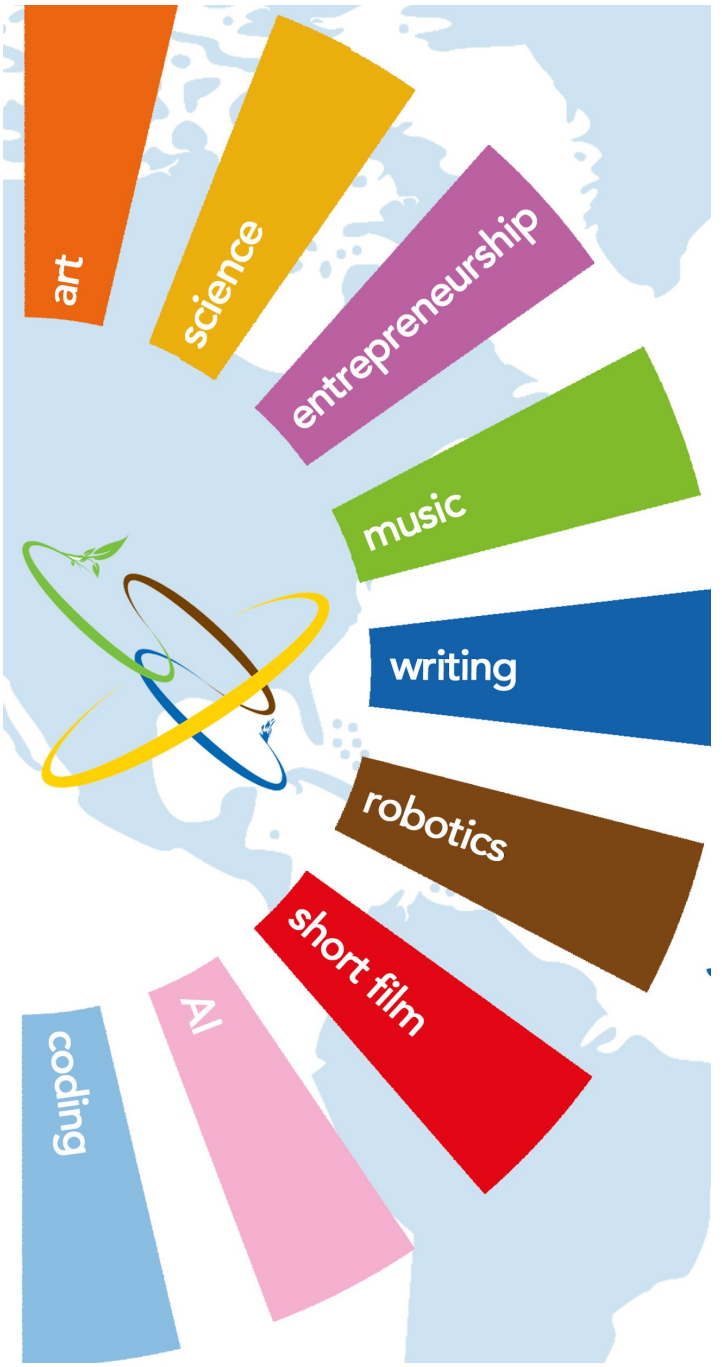
IJHSR International Journal of High School Research

April 2026 | Volume 8 | Issue 8
ijhsr.terrajournals.org
ISSN (Print) 2642-1046
ISSN (Online) 2642-1054

GENIUS OLYMPIAD

Let's build a better future together

International Project Fair focused on Sustainability and Environment for Grades 8-12



- Since 2011
- Hosted around 1400 participants in 2025 from 35+ states 70+ countries
- **Disciplines:** STEM, Coding, Robotics, AI, Speech, Entrepreneurship, Arts, Short Film, Music
- Applications start on December 1
- Application Deadline is March 1
- Finalists are announced by March 25
- Event is usually scheduled 2nd week of June
- Monday – Friday, includes a trip to Niagara Falls
- Hosted by large universities at Upstate New York
- Application Fee is $60/ project
- Participation Fee is $600/ person, w/ room and board
- Open buffet breakfast, lunch, and dinner
- Trip to Niagara Falls and boat tour is included
- **Instagram and Facebook @Geniusolympiad**
- For more information: **GENIUSOlympiad.org**
- Email: **info@geniusolympiad.org**

GENIUS Olympiad is organized by Terra Science and Education, a N.Y. based 501.c.3 non-profit organization dedicated for project-based learning

Marine Biology Research at Bahamas

Unique and exclusive partnership with the Gerace Research Center (GRC) in San Salvador, Bahamas to offer marine biology research opportunities for high school teachers and students.

- Terra has exclusive rights to offer the program to high school teachers and students around world.
- All trips entail extensive snorkeling in Bahamian reefs as well as other scientific and cultural activities.
- Terra will schedule the program with GRC and book the flights from US to the GRC site.
- Fees include travel within the US to Island, lodging, meals, and hotels for transfers, and courses.
- For more information, please visit terraed.org/bahamas.html

Terra is a N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education

Table of Contents

April 2026| Volume 8| Issue 8

Editorial Board

International Journal of High School Research

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Cointegration Analysis of the Performance of Semiconductor Manufacturers by Leveraging the Trend of Big Technology Companies

Alex Taeyeon Shin

St. Paul Academy Daechi, 427, Yeoksam-ro, Gangnam-gu, Seoul, South Korea; shin44649@gmail.com

ABSTRACT: In this study, we evaluate the interrelationship between the performance of major technology companies and semiconductor manufacturers during an unstable economic cycle using Johansen's cointegration test. Recent market instability in the semiconductor industry, stemming from the COVID-19 pandemic and a global trade war, has impacted the semiconductor supply chain across the US, the EU, and Asia. We hypothesize that this environment led to an alignment in the performance between the chipmakers and the technology companies that are their primary consumers. Based on a technical analysis of stock indicators, Johansen's cointegration test was employed to evaluate the extent of the relationship between their respective stock trends. Our findings reveal a strong relationship between the stock trends of technology and semiconductor companies, particularly when these trends are measured by price momentum (the strength or speed of price movement). This suggests that cointegration in the chip market becomes stronger in the short term, primarily influenced by price momentum rather than long-term trading fluctuations or price changes driven by external macroeconomic factors.

KEYWORDS: Systems Software, Languages and Operating Systems, Cointegration Analysis, Technical and Statistics.

Introduction

Semiconductors are a solid material that has a unique electrical conductivity between that of metallic conductors and organic insulators, making them essential for creating integrated circuits and other electronic components.[1] The semiconductor industry is fundamental to driving technological innovation.[2] Devices such as smartphones, computers, electric vehicles, and even household appliances rely on semiconductors for processing power, connectivity, and user interface.

From 2019 to 2024, and continuing to the present, the semiconductor industry has faced significant disruptions due to the COVID-19 pandemic.[3,4] In addition, recent developments in intellectual property protection and geopolitical interventions—such as trade wars—have severely impacted global chip production, leading to shortages, supply chain disruptions, and pricing volatility.[4,5] In particular, the U.S.–China trade war has undermined the stability of the global chip supply chain through export restrictions, sanctions, and efforts by both countries to repatriate manufacturing.[5] This recent economic conflict has prompted economists and policy analysts to reconsider the importance of the relationship between chipmakers and technology companies as chip consumers.[6] Therefore, developing a method to evaluate this interrelationship is essential for understanding the current trend of the chip industry and for forming reliable expectations of the future chip cycle. In this study, we aim to establish a framework for assessing this relationship through the stock trends of semiconductor companies.

In the empirical analysis, three representative chipmakers and three major technology firms are selected to serve as proxies for semiconductor manufacturers and consumers, respectively. The selected chipmakers are Samsung Electronics, SK Hynix, and Taiwan Semiconductor Manufacturing Company, while the consumers are Apple, Lenovo, and NVIDIA. These companies are chosen for their market leadership in the global technology and semiconductor industry, based on regular reports.[7,8] We first analyze their stock trends using technical indicators and then apply the Johansen cointegration test. The technical analysis is applied to the stock markets for visual clarity of price trends and movements, which helps identify trend patterns with entry and exit points.[9] Johansen cointegration is used as the key methodological approach to estimate the extent of the interrelationship among multiple time series variables.[10]

In this study, the selection of stock indicators is critical, as employing a diverse set that captures different dimensions of market behavior allows for more meaningful and reliable results while enhancing the robustness of the analysis. There are several stock indicators, such as trend, momentum, and volatility indicators, to estimate the interrelationship between two industries. Each indicator, however, provides distinct dimensions of market behavior, as shown by long-run co-movements, short-term dynamics, and risk-related fluctuations, respectively. When combined, they complement one another and strengthen the robustness of correlation analysis.

Importantly, we assume that while uncertainty escalates in the semiconductor market for a short period, the stock price movements of big tech companies and chipmakers are expected to be linked. However, this connection would weaken over time due to market resilience and the drive to maintain a balance between supply and demand. This expectation is reasonable if considering the conventional wisdom for the macroeconomic cycle. Raunig and Houari claim that there is no transmission

DOI: 10.36838/IJHSR88.1

mechanism to reveal the complexity and potential negative impact during the shock wave from business uncertainty cycles.[11,12] And the precise linkage between the unstable stock market and macroeconomic aggregate demands has not been well understood. Therefore, our study would provide a pathway to understand the market elasticity in the semiconductor industry for both suppliers and demanders during periods of economic instability.

Methods

In this study, we employed a two-fold approach to investigate the interrelation between the performance of chip manufacturers and major technology companies as their primary consumers. First, we applied technical analysis to assess each company's stock performance, including its price movement and trading volume.[13,14] Second, we applied Johansen's cointegration test to evaluate the strength and extent of the relationships among the variables derived from the technical analysis. The technical analysis helped us identify patterns and trends in stock charts, which differ from fundamental analysis, which focuses on intangible factors such as a company's financial health. Johansen's cointegration test is a robust method for analyzing relationships among multiple time series variables, yielding meaningful insights into long-term relationships. This test has various applications for multivariate analyses such as portfolio management, forecasting, and financial modeling.[10] In this context, a "time series" refers to continuous observations over a specific period of time.

Figure 1 illustrates the flowchart of the research process we applied, which outlines the sequence of the technical analysis and the Johansen cointegration test, with the arrows and boxes indicating the workflow and the sub-processes for each step, respectively. The analysis was performed based on the Python scripts of Google Colab, a free, cloud-based Python notebook environment provided by Google.[15] The scripts are shared in the link.[16]

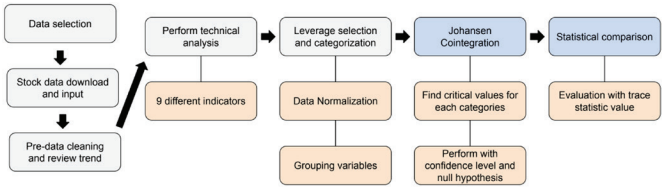


Figure 1: A flowchart of the overall analysis applied in this study. The primary analysis consists of the technical stock analysis and Johansen cointegration test, with a gray and light blue background, respectively. The sub-processes are indicated with an orange background. The arrows indicate the direction of the workflow. The process was irreversible, but parallel for each variable. This flow chart provides an easy guideline of the entire process applied in this study.

The three major semiconductor manufacturing companies for chipmakers and the three leading technology companies as semiconductor consumers are selected for this study.[7,8] The semiconductor manufacturers are Samsung Electronics Co., Ltd (SE: 005930.KS), SK Hynix Inc. (000660.KS), and Taiwan Semiconductor Manufacturing Company Limited (TSM), all known for producing memory and system chips. The tech companies are Apple Inc. (AAPL), Lenovo Group Limited (LNYGY), and NVIDIA Corporation (NVDA). Samsung and SK are listed and traded on the Korea Composite Stock Price Index (KOSPI). In contrast, TSMC, AAPL, LNYGY, and NVDA are traded in American Depositary Receipts (ADRs) on the New York Stock Exchange (NYSE) and the Nasdaq Stock Market (NASDAQ).[8]

Different from the usual economic cycles, the recent instability in the semiconductor sector is supposed to be driven by COVID-19 and geopolitical tension.[3] It seems to be closely linked to strategic competition over emerging technologies like artificial intelligence (AI) and military systems.[4,5] To properly reflect these factors, we analyzed the stock trends of the above companies from January 2, 2019, to December 30, 2024, assuming that this time period captures the dynamics of the global semiconductor conflict.[3] We believe that the relationship between chip manufacturers and chip consumers would become sensitive amid these intensified economic tensions, and it can evaluate the impact of how the performance of the chip-consuming companies influences the market performance of chip manufacturers.

Technical Analysis:

We downloaded the stock trend dataset of our selected companies from Yahoo Finance, which has open, high, low, close, and adjusted close prices.[8] The data was then imported into a Google Colab script for the analysis. During the pre-processing stage, we cleaned the data by removing entries from non-trading days (e.g., holidays and weekends) and periods with anomalous activity caused by extreme volatility or regulatory events. After that, the overall trends were presented as the candlestick chart, as shown in Figure 2. For the technical analysis, we utilized nine different indicators, which are derived from different mathematical models to characterize the various market behaviors. Table 1 describes the features of the indicators applied. The detailed definition and application of the technical indicators are well described in the literature.

Table 1: Brief description of the advantages and disadvantages of the stock indicators applied in this study. From the table (Table 1), one can easily compare the characteristics of the stock indicators for the analysis.

Category	Examples	Advantages	Disadvantages
Trend Indicator (Describing the direction and strength of a trend.)	Simple moving average (SMA), Exponential moving average (EMA), Moving average convergence divergence (MACD)	- Capture long-term price direction and structural co-movements - Filter out short-term noise, useful for equilibrium analysis	- Lagging in nature, less responsive to sudden shocks - May distort short-run correlations
Momentum Indicator (Presenting the speed of price changes)	Relative strength index (RSI), Stochastic Oscillator (SO)	- Measure speed and intensity of price changes - Detect short-term lead–lag relationships and asymmetric responses	- Highly sensitive to noise, risk of spurious correlations - Focused on short-term fluctuations rather than long-term equilibrium
Volatility Indicator (Describing the magnitude of a stock's price movement)	Bollinger Bands	- Reflect the magnitude of price fluctuations - Capture co-volatility during crises or uncertainty	- Lack directional information - Sensitive to extreme values, may overstate correlations

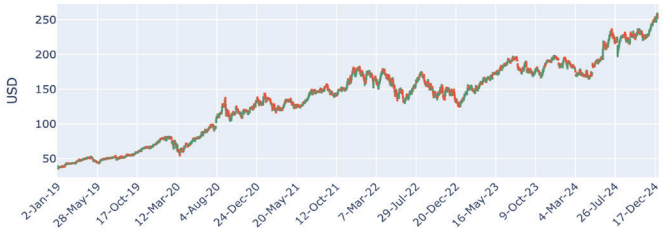


Figure 2: A candlestick chart of Apple Inc. (AAPL) during the selected period, from Jan. 2, 2019, to Dec. 30, 2024. This chart represents the overall price movement of a stock for open, high, low, and close (OHLC). It exhibits the increase in overall prices as the date progresses. Some of the stocks show similar trends, but some show flat behavior.

The 50- and 200-day simple moving average (SMA), and the 12- and 26-day exponential moving averages (EMA), are plotted for Apple's (AAPL) adjusted closing price, as presented in Figure 3. We found that shorter-term averages are more sensitive to changes, compared to longer-term trends for both indicators. The 200-day SMA shows considerable deviation from the actual price movement, while the 50-day SMA generally follows the overall trend. In contrast, both the 12- and 26-day EMAs show close tracking movement against the actual price, but the 26-day EMA shows a slightly late response. This pattern is similarly observed across the price movements of other companies.

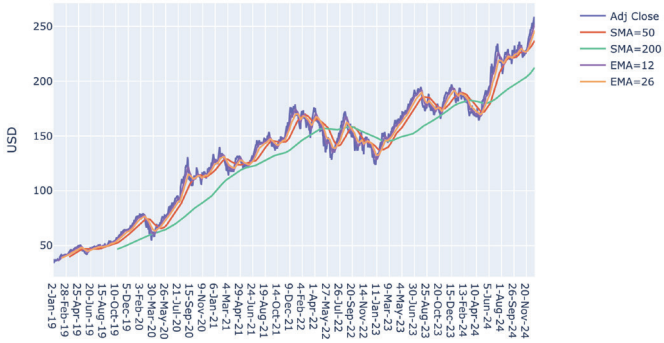


Figure 3: A historical chart of the adjusted close for Apple Inc. (AAPL) during the selected period (from Jan. 2, 2019, to Dec. 30, 2024) with 50- and 200-day simple and 12 and 26-day exponential moving average indicators. Simple and exponential moving averages exhibit upward trends, coinciding with the increase in the adjusted close over the specified time period. However, there is a slight difference between the two moving averages, depending on the calculation method.

The moving average convergence divergence (MACD) and its signal line (MACD-signal) are commonly used to detect trend reversals and to identify potential entry and exit points in the equity market.[15] The range scale defines the MACD is defined by the range scale by subtracting the 26-day EMA from the 12-day EMA, from which a fast-moving line can be obtained. The MACD signal line is a slow-moving line, defined as the 9-day EMA of the MACD line itself. Although the MACD and MACD signal can sometimes be classified as momentum indicators, we treated them as trend indicators in this study. As shown in Figure 4, the signal line moves more smoothly than the MACD line; however, both generally reflect similar strengths in the price movements of AAPL.

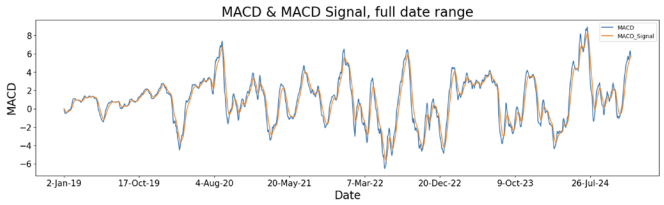


Figure 4: Moving average convergence and divergence (MACD) and MACD-signal indices obtained from the adjusted close of Apple Inc. (AAPL). Both indices exhibit similar strengths in upward and downward movement, but the MACD signal shows a slightly smoother transition. The difference between these two indicators is likely mostly negligible.

As momentum indicators, the relative strength index (RSI) and stochastic oscillator (SO) were applied to measure the magnitude and velocity of change in the adjusted closing price of the stocks. The RSI chart, as shown in Figure 5, measures momentum on a scale of 0 to 100, helping to identify potential trading opportunities. It typically indicates overbought conditions when the value is above 70—suggesting a potential selling opportunity—and oversold conditions when the value is below 30—suggesting a potential buying opportunity.[17] The SO works in a similar way, which implements a scale of 0 to 100 to identify the signals of the possible trading action. It indicates overbought and oversold conditions for above 80 % and below 20 %, respectively. In the SO chart (Figure 6), two indices are shown: %K (STOK) and %D (STOD). %K represents the current adjusted closing value relative to the 14-day range, while %D is a 3-day moving average of %K, resulting in a smoother signal. Figure 6 displays the SO index along with the adjusted close of Apple Inc. (AAPL). Even as AAPL's adjusted close price trends upward over the observed period, the STOK and STOD lines oscillate upward or downward, highlighting potential buying and selling opportunities.

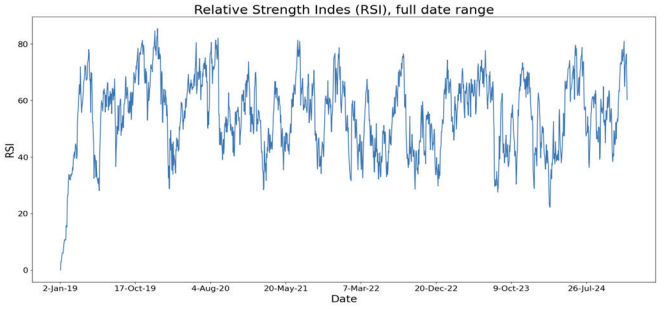


Figure 5: Relative Strength Index (RSI) based on the adjusted close of Apple Inc. (AAPL). The index exhibits both upward and downward movement momentum of the adjusted closing stock price of AAPL. As expected from the mathematical model, the indicator describes the moment of buying and selling trends.

DOI: 10.36838/IJHSR88.1

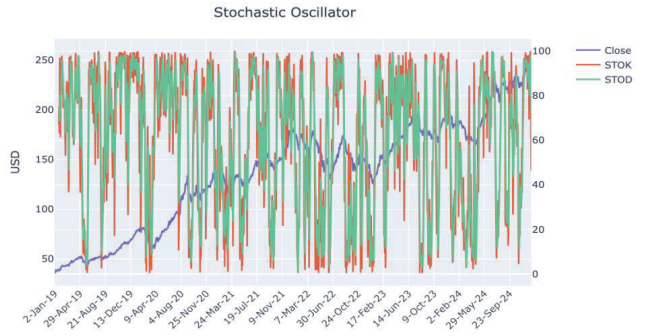


Figure 6: The stochastic oscillator (SO) index, along with the adjusted closing price of Apple Inc. (AAPL), is indicated by STOK and STOD. (%) K presents a 14-day moving average of the closing price, and (%) D indicates a smooth line of K over 3 days. The left y-axis indicates the adjusted closing price, and the right axis indicates the range of 0-100%. The difference between STOK and STOD is hardly noticeable, but it seems that both show very similar trends.

As a volatility indicator, the Bollinger Bands (BB) are plotted against AAPL's adjusted closing price to illustrate price volatility and identify potential trading opportunities, as shown in Figure 7. In this context, volatility refers to the degree of price fluctuations in the stock market and serves as a measure of price stability. The middle line of the BB is the 20-day simple moving average (SMA), while the upper and lower bands are calculated as two standard deviations above and below the SMA, respectively. As shown in Figure 7, the width of the Bollinger Bands expands alongside increases in AAPL's adjusted closing price, reflecting greater volatility.

We were able to build a comprehensive picture of each company's unique stock behavior using this set of technical indicators, as described above. This holistic view was essential for investigating the market relationships at the core of our study. And the individual outcomes were applied for the Johansen co-integration test.

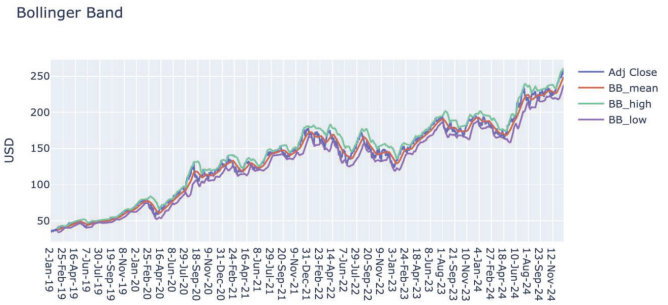


Figure 7: Bollinger band (BB) lines with the adjusted close price of the AAPL chart. BB lines are indicated by three lines: middle (mean), high, and low lines, which are acquired by a 20-day simple moving average (SMA) and two standard deviations for the above or below value. The three individual trends of the BB indicator are moving simultaneously along with the date proceeds, and the different exist, considering the statistical meaning.

Johansen Cointegration:

Before running our main cointegration comparison, we had to pre-process the data obtained from the technical analysis. It is because the companies' stocks are traded in different currencies. We normalized all the values with the mean of the statistics to ensure a fair, apples-to-apples comparison, as shown in Figure 8. As described in Table 2, we paired each of the three major chip-consuming companies with each of the three chip manufacturers in one-on-one comparison. This pairing was effective for assessing the specific influence that chip makers have on the market movement of the companies that use their chips. Finally, the Johansen cointegration test was performed.

In the cointegration test, the first step was to calculate the critical values under the null hypothesis of zero cointegration relationship (r = 0). Then, we sequentially computed the values for $r > 1$ and $r > 2$. Here, "r" represents the rank, indicating the number of cointegration vectors under the null hypothesis.[9] The critical values serve as the thresholds for determining whether different non-stationary time series data can be statistically related in a long-run equilibrium.

Next, we calculated the trace statistics for each pair as shown in Table 2, based on Equation 1. This trace statistic value is the number that we can compare against the critical values, which provides a mathematical proof needed to determine how many cointegration relationships (r) exist among variables.[9]

$$Tace\ Statistics = -T \sum_{i=r+1}^{k} \ln(1-\lambda_i) \qquad (1)$$

Here, T is the number of observations, K is the number of variables, and λ_i is the i-th eigenvalue from the decomposition.

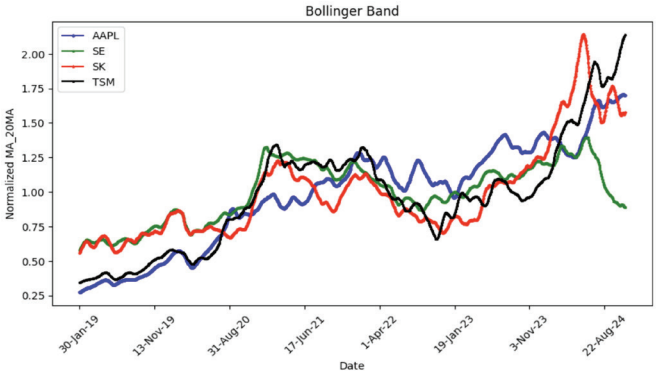


Figure 8: The normalized Bollinger Bands (BB) indices of Apple, Inc. (AAPL), Samsung Electronics (SE), and Taiwan Semiconductor Manufacturing Company Limited (TSM) are based on the indication by STOD (D%). Each curve shows a different behavior, even though they gradually increase along with the progression of the data.

Table 2: The pairs of the chip makers and consumers for the Johansen cointegration test. The A, B, and C in the parentheses define the individual pair. This is the simple combination used for this study, only to compare each big tech company's performance.

Semiconductor Consumer Big Tech	AAPL (A)	LNVGY (B)	NVDA (C)
Semiconductor Major Manufacturer	SE (A1)	SE (B1)	SE (C1)
	SK (A3)	SK (B2)	SK (C2)
	TSM (A3)	TSM (B3)	TSM (C3)

Results and Discussion

The critical values for the Johansen cointegration test are not derived from a simple analytical formula. Still, they are obtained from simulated distributions corresponding to specific confidence levels, such as 99%, 95%, and 90% (i.e., significance level of 1%, 5%, and 10%, respectively). For our analysis, we used Python's *Statsmodels* library, which provides these simulated values for both the trace and maximum eigenvalue tests in the asymptotic distribution approach.[9,10] We utilized this library to produce the critical values for our specific hypothesis test for r = 0, ≤ 1, and ≤ 2, which are presented in Table 3 for each confidence level. These confidence values were initially acquired during the Johansen cointegration analysis. Each confidence value numerically exhibits the maximum cointegration that can be rejected for the level of the null hypothesis, without considering eigenvalues for the numerical calculation of the cointegration. Namely, these are used as the minimum values for cointegration for each pair.

Table 3: The critical values for the cointegration confidence level and the null hypothesis rank (*r*). The numbers indicate the minimum values for each category, specifically the values required for confidence optimization. The values decrease as the confidence level increases.

Confidence Level	Column 1 (r =0)	Column 2 (r <= 1)	Column 3 (r <= 2)
90%	44.49	47.85	54.68
95%	27.07	29.80	35.46
99%	13.43	15.49	19.93

The trace statistics were also calculated for each pair in Table 2 with the cointegration rank via Python's Statsmodels package. If the trace statistic exceeds the critical value at a given rank (*r*)—that is, if *trace statistic* > *critical value*—the current null hypothesis can be rejected, and the test proceeds to the next rank. For example, in the case of $r = 0$ in Column 1, if the trace statistic exceeds 44.49 at a 90% confidence level, the null hypothesis of no cointegration is rejected, and we move on to test for $r < 1$. This implies that there may be at most one cointegration relationship, indicating that the variable—such as a stock indicator—exhibits more than zero cointegration relationship between the chip makers and consumers. We applied this procedure to all pairs listed in Table 2 and present the resulting trace statistics in Table 4. To make the table easy to read, we color-coded the results: A green background in the table indicates no cointegration relationship. In contrast, an orange background indicates the presence of at least one cointegration between the variables. The values in Table 4 present the trace statistics from the Johansen cointegration test for each pair of stock indicators in Table 2. Since the eigenvalues (specific scalar approach) are not considered in this study, the trace statistic from the cointegration test is used to determine whether the null hypothesis is rejected at each confidence level. Therefore, the trace statistics are important for examining the cointegration. As marked with a colored background in each cell, the moving indicators show no integration, while the momentum indicators exhibit strong cointegration.

We also created a 3D plot showing the trace statistics values for each pair alongside their corresponding stock indicators, which makes it easy to identify the high trace statistic values. The high values are prominent because they indicate the presence of at least one cointegration relationship between the indicator variables.

The results of our study suggest that during periods of broad economic instability, the stock trends of semiconductor manufacturers and the major technology companies that consume their chips may become closely aligned. At this time, tech companies manage their production to minimize losses, so that this decision directly impacts chip demand. This tendency is likely to become predominant with a dynamic intensified by the recent turmoil from the COVID-19 pandemic and rising trade protectionism. As addressed in the literature,[11,12] although the business model has not been well developed for full capturing of the volatility dynamics of the financial and economic cycle, empirical and existing theories have shown the stock's sensitivity and co-movement to economic fluctuation across various economies.[18]

For trend analysis across multiple variables, time-relevant features need to be considered to improve analysis performance. However, in this study, we only examined the defined cointegration among commercial market equities over a specific time period. This may be why our results show a distinct difference between momentum and trend indicators. We suggest that time-series analysis should be advanced in correlation or cointegration analysis, or that such analysis be performed for stronger quantitative evaluation. This will be proposed for our next project. In addition, our model does not reflect intangible macroeconomic parameters such as interest rates and the Korean won-to-US dollar exchange rate; as a result, the cointegration results likely further demonstrate the impact of these unobserved features rather than pure firm-level interactions. Indeed, a small sampling (only six companies: three semiconductor manufacturers and three semiconductor consumers) would limit the confidence in the more relevant analysis.

For the time-series analysis, our technical approach did not include trajectory evaluation in a time period, which is typically used to determine whether the trend is stationary or non-stationary. The fact of this evaluation is beneficial for systematically estimating the relation between the trends of two variables over time. This evaluation will be able to effectively discriminate between a pair of companies with similar trend fluctuations among multiple companies. In addition, asymmetric market analysis was not included in our study, so our results may not apply to other specific time periods or semiconductor economic cycles. Such analysis is called volatility-spillover analysis. We will also include these techniques for future analysis and believe that this proposed approach will provide further details on the correlation and cointegration analysis.

DOI: 10.36838/IJHSR88.1

Table 4: The trace statistic values from the Johansen cointegration test for each pair of stock indicators are presented in Table 2. The moving indicators show no integration, while the momentum indicators exhibit strong cointegration.

Big Tech	Manufacturer	50SMA	12EMA	MACD_Signal	RSI	SO	BB
APPL	SE	25.68	17.40	207.14	104.57	179.19	22.66
	SK	7.74	7.59	176.82	86.90	159.11	6.37
	TSM	5.86	4.59	126.81	73.99	151.08	5.25
LNVGY	SE	26.91	24.79	213.59	114.63	164.52	24.40
	SK	12.04	13.42	179.91	84.82	158.71	10.36
	TSM	7.28	5.45	142.12	76.46	149.48	8.77
NVDA	SE	37.89	29.31	218.40	104.77	171.02	37.32
	SK	23.63	16.74	199.66	82.88	158.53	23.53
	TSM	6.18	7.02	90.49	75.86	151.30	4.67

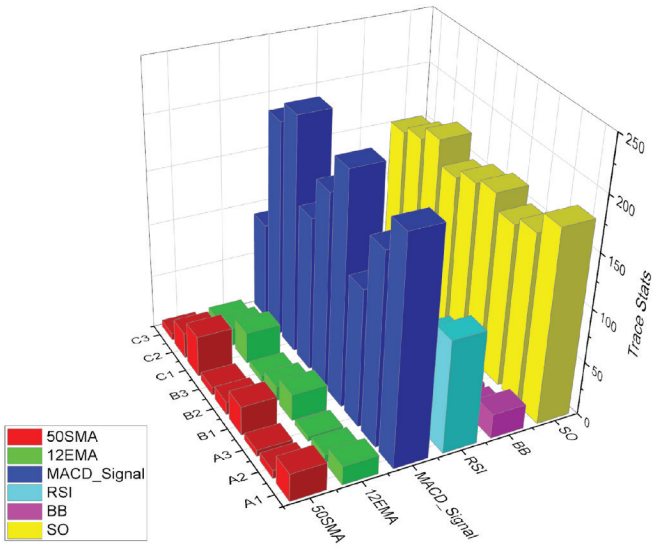


Figure 9: 3D plot of the trace statistics for each pair in Table 2 from the Johansen cointegration test. This plot provides an easy-to-read representation of the level of cointegration in terms of the trace statistics.

Conclusion

In this study, we evaluated chip makers' performance by leveraging the stock trends of their key consumers. Based on the technical analysis, we applied the Johansen cointegration test to assess the relationships between these groups quantitatively. Although the data selection is limited to only six companies in total, our technical approach highlights the distinct interdependencies in the semiconductor economy. The finding can be beneficial for understanding the market dynamics and the complex nature of global trade and supply chain instability in the short term.

It is observed that momentum indicators—such as the Relative Strength Index (RSI) and Stochastic Oscillator (SO)—are most effective at revealing the strong relationship between chipmakers and consumers. Different from the trend indicators or volatility indicators, the moment indicator shows a close relationship by measuring the speed of price changes. However, the MACD indicator is an interesting exception. Even though we classified the MACD as a trend indicator, it did show a solid connection. It would be because its formula is based on the incorporation of the elements of both trend and momentum. Likely, the MACD seems to capture not only the market direction but also the strength of price movements.

Based on our findings, the indicators that measure price momentum (strength and speed) are more suitable for analyzing the semiconductor market, compared to the indicators that track general trends or price fluctuations. In addition, the critical cointegration cannot be considered for the movement of the trends in the long-term direction. Still, it will be more prominent when reacting to market pressures with similar force and speed in short-term volatile times. However, in our technical analysis, the asymmetric market evaluation, volatility-spillover study, and correlation related to stationary or non-stationary movement are not included due to technical complexity. Such analysis would be beneficial for further effective discrimination of the individual markets pairing and performance evaluation, regarding macroeconomic parameters. Nevertheless, we believe that our findings can help investors and economic analysts make more informed decisions by highlighting the strong underlying connection between technology and semiconductor firms, especially in rapidly changing markets.

Acknowledgments

I am grateful to my family for their unwavering support. I could not have completed this project without their unconditional support. The author would like to thank Dr. Sung Z-H for his initial motivation and invaluable guidance throughout this project.

References

1. Singh, M; Sargent, Jr. J. F.; Sutter, K. M. Semiconductors and the Semiconductor Industry. Congressional Research Service (CRS) R47508. P.L.116-283; P.L.117-167, 04-19-2023.
2. Yeboah, L.; Oppong, P.; Abdul Malik, A.; Acheampong, P.; Morgan, J.; Addo, R.; Henyo, B. W. Exploring Innovations, Sustainability and Future Opportunities in Semiconductor Technologies. Int J Adv Nano Comput Anal, Vol 3 No 2, 2024 doi: 10.20944/preprints202409.1307.v3.
3. Haramboure, A.; Lalanne, G.; Schwellnus, C.; Guilhoto, J.; Vulnerabilities in the semiconductor supply chain OECD Science, Technology and Industry Working Papers, 19 June 2023. https://dx.doi.org/10.1787/6bed616f-en.
4. Ivanova, M.; Ivanov, D. Responding to the ripple effect from systemic disruptions: empirical evidence from the semiconductor shortage during Covid-19. Modern Supply Chain Research and Application, 6 (4): 354-375, (2024). https://doi.org/10.1108/MSCRA-03-2024-0011.
5. Alexander, P. The price of Semiconductors and the Macroeconomy: A real Business Cycle Approach, Economics, June 23, 2025.
6. Bibi, J. The US-CHINA Semiconductor War: How it impacts South Korea's tech industry, International Journal of Social Science Bulletin, Vol 3, Issue 6, 3007-1909, 25 June 2025.

7. Companies Market Cap, "Largest semiconductor companies by market cap" https://companiesmarketcap.com/semiconductors/largest-semiconductor-companies-by-market-cap/
8. Yahoo Finance, "20 Largest Semiconductor Companies in the World", https://finance.yahoo.com.
9. Dwyer, G. P. The Johansen Test for Cointegration. v516, White Paper, April 2015.
10. Poh, C. W.; Tan, R. Performance of Johansen's Cointegration Test. In East Asian Economic Issues (Applied Economics Research Series), East Asian Economic Issues, pp. 402-414 (1997). https://doi.org/10.1142/9789812819376_0029.
11. Burkhard, R.; Scharler, J. Stock market volatility and the business cycle. Monetary Policy and the Economy 2.10 (2010): 54-63.
12. Houari, O.; Uncertainty shocks and business cycles in the US: New insights from the last three decades. Economic Modelling, Volume 109, 2022, 105762. https://doi.org/10.1016/j.econmod.2022.105762
13. Petrusheva, N.; Jordanoski, I. Comparative Analysis Between the Fundamental and Technical Analysis of Stocks. Journal of Process Management – New Technologies, International Vol. 4, No.2, 2016
14. Pramudya, R.; Ichsani, S. Efficiency of Technical Analysis for the Stock Trading. Pramudya and Ichsani / International Journal of Finance & Banking Studies, Vol 9 No 1, 2020. https://doi.org/10.20525/ijfbs.v9i1.666
15. Cold Colab, https://colab.research.google.com/
16. https://drive.google.com/file/d/1aDCPhVQ8z8Zn_5fTJbkhyR3uNZiuem6m/view?usp=sharing
https://drive.google.com/file/d/1uflWeGezJocnHkDdZ3Vz4KdKGgmlf9-v/view?usp=sharing
17. Kofi Nti, I.; Adekoya, A. F.; Weyori, B. A. A systematic review of fundamental and technical analysis of stock market predictions" Artificial Intelligence Review (2020) 53:3007–3057. https://doi.org/10.1007/s10462-019-09754-z
18. Qu, Z. A Review about the Business Cycle and Stock Market, Advances in Economics Management and Political Sciences 124(1):234-238, November 2024, DOI:10.54254/2754-1169/2024.17720

Author

Taeyeon Shin is a sophomore at Saint Paul Daechi Academy, interested in business with plans to pursue a major in business administration. He hopes to build his career as a business leader in the field of international commerce.

DOI: 10.36838/IJHSR88.1

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Pollination Behavior in Bees: What *Wolbachia* Studies in *Drosophila* Can Teach Us

Sean Seohyun Park

Burnaby North Secondary School, 751 Hammarskjold Drive, Burnaby, BC, V5B4A1, Canada; sean090309@gmail.com

ABSTRACT: *Wolbachia* is an endosymbiotic bacterial species that is transmitted through the female germline. *Wolbachia* use several strategies to enhance the reproductive success of their infected hosts. Recent evidence suggests that *Wolbachia* proteins in the brain manipulate host behavior to increase their reproduction. In the fruit fly *Drosophila Melanogaster, Wolbachia* colonize brain regions that are associated with mating and reproduction behaviors. This raises the question of whether *Wolbachia* alter behaviors in other insect species. Insects infected by *Wolbachia* include a number of *Anthophila* bee species, which are important pollinators of flowers. In these bees, the same brain regions that *Wolbachia* colonizes in fruit flies are involved in more complex behavior beyond reproduction, including pollination. Direct evidence of *Wolbachia* affecting the brains of *Anthophila* bees has not yet been established, but growing evidence of structural similarities (e.g., brain) between *Drosophila* and bees suggests that *Wolbachia* may influence bee behavior as well. The review of existing literature suggests that *Wolbachia*'s effect on ecological decision-making in pollinators can be further investigated, as *Wolbachia*'s effect on pollinators remains underinvestigated.

KEYWORDS: Cellular and Molecular Biology, Neuromicrobiology, Brain Homology, Pollination, Food Chain.

Introduction

Wolbachia are highly successful bacterial endosymbionts that infect a vast array of arthropods and filarial nematodes. Their remarkable evolutionary success stems from their ability to manipulate host behavior and reproduction. On a molecular and cellular level, *Wolbachia* interact with a broad range of host-derived cellular components to adapt and thrive in diverse cell types and environments.[1]

Recent studies suggest that *Wolbachia* proteins in the brains of infected fruit flies manipulate host behavior to increase their reproduction. In the fruit fly *Drosophila Melanogaster*, *Wolbachia* colonizes brain regions that are associated with mating and reproduction behaviors.[2] However, studies have mainly focused on fruit flies and other insects through understanding cytoplasmic incompatibility, male killing, parthenogenesis, and feminization.[1] Little is known about whether *Wolbachia* alter the behaviors of bees that play significant roles in pollination and the ecosystem. However, as previous research has mainly focused on fruitflies, little is known about whether *Wolbachia* alter behaviors in other insect species that play significant roles in pollination and the ecosystem.

Given the similarities in brain structure between fruit flies and bees,[3] I hypothesize that the effects of *Wolbachia* on host behavior may also extend to bees. The study of the effects of *Wolbachia* infection on *Anthophila* bees is particularly important because it can help elucidate manipulations of pollination, specifically floral selection preferences and pollen-collecting patterns. Additionally, such modifications can disturb the food chain, triggering a chain reaction that affects animals across the population. Extension of research on the effects of *Wolbachia* to include *Anthophila* bees provides insights into how the consequences of endosymbionts on pollinators may influence ecosystem stability.

Building on existing research, this study examines whether *Wolbachia* infection influences pollination behavior in bees, specifically their floral selection and pollen-collecting behaviors. Since bees play a key part in ecosystem functioning and flower reproduction, this study has the potential to reveal unexplored mechanisms through which symbiotic bacteria impact pollination and biodiversity.

Discussion

Wolbachia Infection and Host Colonization:

Wolbachia are bacterial endosymbionts of arthropods and some filarial nematodes. It is estimated to infect 40-60% of all insect populations, such as mites, isopods, and worms.[1,4] *Wolbachia* are vertically transmitted through infected female oocytes. *Wolbachia* employ several well-characterized strategies to ensure their reproductive success in mixed populations of infected and uninfected hosts. More recently, an intriguing report suggests that *Wolbachia* might alter host behavior to increase mating by infected hosts.

Although *Wolbachia* are vertically transmitted through infected female oocytes, they also colonize somatic tissues. The full extent of *Wolbachia*'s cellular localization remains incompletely mapped.[5] A study identified distinct patterns of oocyte infection, primarily localized at the posterior pole of the germline, suggesting targeted infection of female reproductive structures.[6] *Wolbachia* have also been found in the brain, neural cells, salivary glands, and excretory cells of diverse hosts.[1] Specifically in *Drosophila* fruit flies, *Wolbachia* has been known to induce tissue-specific behaviors through the points at which it infects (Figure 1).

DOI: 10.36838/IJHSR88.8

Figure 1: *Wolbachia* (marked in green) is concentrated along the female germline and the brain of a *Drosophila Melanogaster*.

Wolbachia's Effects on Host:

Effects of *Wolbachia* on *Drosophila* behavior have been extensively researched, particularly through studies of its localization in the brain. Albertson *et al.* mapped *Wolbachia* distribution in the brain and found that the bacterium interferes with cells involved in neural signaling.[7] Using 3D imaging, they showed that *Wolbachia* is not evenly distributed throughout the brain, but instead is concentrated in specific regions. High densities of *Wolbachia* were observed in the antennal lobes (olfactory-guided behaviors), mushroom bodies (learning and memory), optic lobes (vision-based behaviors), and the central complex and suboesophageal ganglion (locomotor and feeding behaviors), each corresponding to distinct behavioral functions. *Wolbachia* was particularly associated with ganglion cells, supporting and nourishing the neural environment of the brain. The particular strains of *Wolbachia* and the species of *Drosophila* had subtle influences on infection patterns. Furthermore, female *Drosophila* exhibited slightly higher *Wolbachia* densities in the brain than males.

Wolbachia has been shown to influence *Drosophila melanogaster*, ameliorating learning and memory positively. Studies by Chen *et al.* indicated that *Wolbachia* infection in the brain improves both learning ability and memory performance.[8] They further suggested that these cognitive augments may result from *Wolbachia*-induced changes in host gene expression within neural tissues.

An important consideration in understanding *Wolbachia*-induced behavioral effects across host taxa is the diversity of *Wolbachia* strains. In *Drosophila*, multiple closely related strains, such as wMel, wMelPop, and wRi, have been shown to vary in tissue distribution and behavioral consequences, including effects on sleep, learning, and reproductive behavior. [9] Notably, even within a single host species, distinct *Wolbachia* strains can produce different phenotypes, accentuating the strain-specific nature of *Wolbachia*–host interactions. Likewise, *Anthophila* bees are also naturally infected by diverse *Wolbachia* strains.[10] Given that different *Wolbachia* strains can exert distinct effects on host behavior, as demonstrated in multiple insect systems, it is likely that any influence of *Wolbachia* on pollination behavior in *Anthophila* would depend strongly on the specific strain involved.

Sleep behaviors in *Drosophila* are affected by *Wolbachia* upon infection. Bi and Wang reported that *Wolbachia* can influence the sleep patterns of *Drosophila melanogaster* and *Drosophila simulans*.[2] Using an advanced *Drosophila* monitoring system, they evaluated the activity of the flies and found that *Wolbachia*-infected individuals manifested decreased daytime activity and increased sleep duration.

Wolbachia has also been shown to manipulate the reproductive behaviors of *Drosophila*. Warecki *et al.* demonstrated that *Wolbachia* localizes in brain regions critical for sensory perception and decision-making in *Drosophila melanogaster*, resulting in alterations in reproductive behavior and promotion to promiscuity.[9] These effects were linked to disruption of glutamatergic signaling, as *Wolbachia* infection reduced host metabotropic glutamate receptor (mGluR) abundance, implicating synaptic transmission as a molecular target of behavioral modulation.

Glutamate signaling is a molecular target that *Wolbachia* could influence pollination behavior, as glutamate receptors hugely contribute to sensory integration, decision-making, and motor coordination in insects. In *Drosophila*, *Wolbachia*-mediated reduction of the metabotropic glutamate receptor mGluR has been shown to alter mating behavior by modifying neural signaling efficiency, demonstrating that *Wolbachia* can directly affect neurotransmitter receptor abundance and synaptic function. In pollinating insects, glutamatergic pathways are essential to processing various sensory inputs, including visual, olfactory, and mechanosensory cues, that guide flower recognition, landing, and handling. Consequently, *Wolbachia*-induced modulation of glutamatergic signaling influences foraging efficiency, floral choice, or behavioral persistence on flowers.

The study also identified cellular mechanisms associated with neural transmission. Protein analysis of infected larval brains revealed changes in host proteins and transcription factors, all of which are associated with sensory processing. Additionally, video recordings further showed that infected females had significantly higher copulation rates, despite no difference in mating initiation time. Hybrid mating experiments supported these findings, revealing that infected females were more likely to engage in cross-species mating with *D. simulans* males, resulting in hybrid offspring. Beyond strain identity, an additional approach to assessing the potential for *Wolbachia*-induced behavioral modulation is to examine whether *Wolbachia* strains infecting *Anthophila* have gene families previously identified as candidate effectors in other insect hosts. Comparative genomic analyses have identified ankyrin repeat-containing proteins and Type IV secretion system components in *Wolbachia*, which functions as effector proteins involved in host manipulation.[1,5,11]

In Drosophila, several of these effector classes have been associated with neural, reproductive, and behavioral phenotypes, suggesting that their presence reflects a latent capacity for host behavioral modulation.[9] Although the behavioral roles of these effectors have not been directly characterized in *Anthophila*, molecular surveys indicate that bee-infecting *Wolbachia* strains possess homologous effector gene families.[10] This shared effector repertoire raises the possibility that *Wolbachia* infecting bees retain the genomic potential to influence host neural processes relevant to pollination behavior. Importantly, whether these effectors are expressed in neural tissues or produce behavioral consequences in bees remains an open

question, underscoring a key direction for future functional and comparative genomic studies.

Neurobiological Links Between Drosophila and Anthophila Brain:

In *Anthophila* bees, strong homology exists with *Drosophila* brains.[12] Within the genus of *Drosophila*, multiple species have been independently leveraged to study *Wolbachia*–host interactions. While *Drosophila melanogaster* has been the primary model in many neurobiological and behavioral studies, other *Drosophila* species, such as *Drosophila yakuba*, *Drosophila sechellia*, *Drosophila ananassae*, and *Drosophila recens*, have been central to research on *Wolbachia* transmission dynamics, host manipulation, and cellular interactions.[1,2,4,7] The hypothesis that *Drosophila*-based findings can inform *Anthophila* behavior, therefore, rests not on a single species comparison, but on convergent evidence across multiple *Drosophila* systems, supported by conserved brain architectures involved in sensory integration and behavioral modulation (Figure 2).

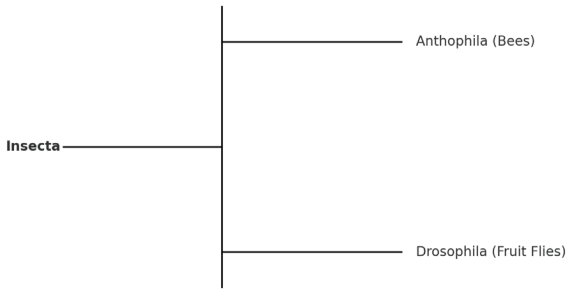


Figure 2: The phylogeny shows the divergence of insect species, such as *Anthophila* bees, *Drosophila* fruit flies, and additional insect species known to host Wolbachia. In several of the taxa shown, particularly *Drosophila* species and Aedes aegypti, *Wolbachia* has been detected in neural tissues and/or demonstrated to alter host behavior, including sensory processing, activity, and reproductive behaviors. In contrast, corresponding neurobiological and behavioral effects of *Wolbachia* in *Anthophila* bees remain poorly characterized. This phylogenetic context highlights that *Wolbachia*-mediated behavioral modulation occurs across diverse insect lineages and provides a framework for evaluating whether similar interactions may exist in bees.

The homology between the bee and fly brain structures has been established through comparative morphological studies, which revealed evolutionarily conserved features across the two taxa. Strausfeld demonstrated that despite vast differences in behavior and ecological niches, insects share conserved morphology in their central nervous systems.[4] Strausfeld and Hirth showed that specific neural circuits, such as those in the mushroom bodies and central complex, are conserved across insect lineages.[12] These shared circuits are crucial for behavioral plasticity, decision-making, and sensory integration, all of which may be modulated by *Wolbachia*.

Comparative studies of insect neuroanatomy have highlighted the organization of the *Drosophila* brain and its functional parallels in other species such as *Apis mellifera*. The *Drosophila* brain contains ~100,000 neurons organized into functional neuropils,[7] including the mushroom bodies and antennal lobes, which are involved in olfactory processing and learning.[13] These structures are critical for modulating behavioral responses to stimuli, such as courtship and foraging. The central complex, another conserved brain region, contributes to motor coordination and spatial orientation.[14] In *Apis mellifera*, a member of the *Anthophila* clade, the brain also contains well-developed mushroom bodies, which are associated with olfactory learning and memory.[15] The antennal lobes process floral odors and pheromones (Galizia & Rössler, 2010), while the optic lobes support visual navigation during foraging. Similarly, the central complex contributes to orientation and locomotion (Figure 3).[14]

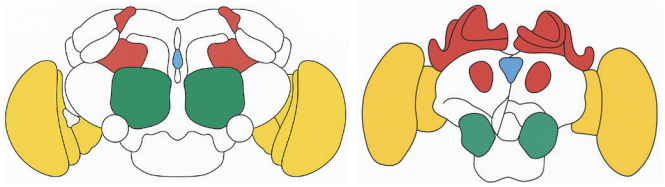


Figure 3: Antennal lobes (marked in green), mushroom bodies (marked in red), optic lobes (marked in yellow), and the central complex (marked in blue) are comparatively analyzed between the two species. The brain on the left represents that of *Drosophila melanogaster*, while the one on the right depicts the brain of an *Anthophila* bee. The general locations of each brain region are similar. The antennal lobes are localized in the lower center of the brain, the mushroom bodies are situated at the top, the optic lobes are positioned along the right and left sides, and the central complex lies at the center of the brain.

Anthophila and *Drosophila* both rely heavily on olfactory and gustatory systems for resource detection, including food and mate location. In these groups, chemosensory function is supported by odorant-binding proteins (OBPs), which enable odorant receptors to solubilize and transport volatile compounds to sensory neurons. In *Apis mellifera*, genomic analyses have identified dozens of OBP genes that are highly expressed in antennal and mouthpart tissues, reflecting strong functional capability in olfaction and taste.[16] Although direct one-to-one orthology between individual *Drosophila* OBPs (such as Obp99b) and bee OBPs has not been fully resolved, OBPs across insects share conserved structure and biochemical roles, suggesting functional conservation despite the divergence.

Given that Warecki *et al.* demonstrated *Wolbachia*-mediated upregulation of specific OBPs in *Drosophila*, including Obp99b, which is expressed in chemosensory tissues and linked to behavior, it is plausible that *Wolbachia* could similarly alter OBP expression in *Anthophila*.[9] Such modulation could occur through conserved regulatory pathways acting on OBPs broadly rather than on a single strict ortholog, potentially influencing olfactory-driven behaviors such as floral preference, foraging efficiency, or pollen collection. However, comparative analyses of OBP sequence similarity, tissue-specific expression, and *Wolbachia* responsiveness in bees remain limited; future molecular and transcriptomic investigations would help support this claim.

Furthermore, both taxa exhibit flexible foraging strategies mediated by neural plasticity in the mushroom bodies. Giurfa demonstrated that bees can learn complex patterns and odors associated with floral resources, dependent on mushroom body function.[11] In *Drosophila*, comparable learning processes have been linked to dopaminergic circuits modulated by environmental cues.[17] Therefore, if *Wolbachia* alters neural processing in flies, as shown by Warecki *et al.*,[9] similar mechanisms may also be present in bees.

DOI: 10.36838/IJHSR88.8

Both *Anthophila* and *Drosophila* are known to be naturally infected with *Wolbachia*. Many studies confirmed natural infections in *Drosophila melanogaster*, with *Wolbachia* localized in specific brain regions. Bee species within *Anthophila* are also frequently infected, although evidence for *Wolbachia* localization in bee brains remains limited. Gerth *et al.*, for example, conducted a molecular screening of multiple *Anthophila* families and found *Wolbachia* infection in the majority.[10] Their analysis revealed substantial variation in infection prevalence across *Anthophila* species and in the *Wolbachia* strains detected.

Insights from the literature on *Wolbachia*'s ability to manipulate the *Drosophila* brain provide new perspectives, particularly for investigating species with more complex behaviors. By infecting brain regions involved in sensory perception and decision-making, *Wolbachia* alters the reproductive behavior of its hosts. However, current studies have primarily focused on *Wolbachia*-targeted regions in *Drosophila* fruit flies. These brain regions are limited to personal behaviors, including sleep, learning, memory, and reproduction.

Pollination Biology and Behavior of Bees:

In contrast to fruit flies, the brain regions of bees are not only linked to individual behaviors, but are also directly stimulated during pollination. While *Drosophila* mating relies on complex neural circuitry, pollination behavior in bees requires learning, navigation, and multimodal sensory processing across repeated foraging events and broader ecological contexts, making it more complex than mating. Visual cues, scents, learning, and memory form the foundation of pollination in bees, with distinct brain regions responsible for these functions. Visual cues, particularly floral color and pattern, strongly influence foraging preferences in *Anthophila* bees. With vision tuned to ultraviolet, blue, and green light, bees are able to perceive intricate floral cues hidden from human eyes.[18] UV nectar guides, such as radial patterns or petal color contrasts, direct bees to floral rewards and increase foraging efficiency.[19] In *Bombus terrestris* and *Apis mellifera*, studies have shown a marked preference for blue and violet flowers, which typically provide higher nectar rewards in natural settings.[20] Floral symmetry and shape, such as radial symmetry and shallow corollas, also affect the speed and accuracy of landing and handling.[21] Together, these visual traits not only attract pollinators but also facilitate rapid flower recognition and efficient reward retrieval, thereby optimizing the energetic return of each foraging trip.

Floral scent acts as a powerful olfactory cue shaping both initial attraction and long-term foraging behavior in bees. The antennal lobes of the bee brain are specialized for detecting volatile organic compounds (VOCs) emitted by flowers, many of which are species-specific.[22] These odors are readily learned and remembered, allowing bees to form associative memories that guide future foraging decisions.[23] In stingless bees such as *Trigona fulviventris*, foragers display strong scent fidelity, returning to flowers that emit familiar VOCs after just one successful visit.[24] Honeybee foragers can also detect scent trails left by conspecifics, using this social information to locate productive floral patches.[25]

Nectar and pollen reward characteristics are central to floral choice. Bees evaluate not only the presence, but also the quality of floral rewards. Studies confirm that they prefer flowers with higher nectar sugar concentration and greater pollen protein content, both of which directly impact colony nutrition.[26] In *Apis mellifera*, and more broadly across *Hymenoptera*, foragers adjust visitation rates according to changes in nectar volume and concentration, reflecting a dynamic assessment of floral profitability.[27] This behavior aligns with optimal foraging theory, which predicts that pollinators maximize net energy gain per unit time. Furthermore, bees can detect and remember floral traits associated with rich rewards, reinforcing constancy to high-quality species.[28]

Handling time and extraction efficiency further shape floral preferences. Morphologically complex flowers, such as those with long or narrow corollas, require more time and skill to handle, discouraging inexperienced or time-constrained foragers.[29] Bumblebees and honeybees often favor flowers that allow rapid nectar access with minimal manipulation, especially when simpler alternatives are available in the same patch.[30] Over time, experienced bees reduce handling time through motor learning, yet innate or early-learned preferences still bias them toward simpler floral structures.[31] This trade-off explains why certain flowers, despite offering abundant rewards, receive few visits: the mechanical costs outweigh the benefits. Thus, handling efficiency acts not only as a foraging constraint but also as a selective filter in plant–pollinator interactions.

The mushroom bodies, responsible for learning and memory, underlie the complex foraging strategies of *Anthophila* bees, particularly flower constancy. These brain structures integrate multisensory experiences and guide learned behaviors.[15] Through repeated exposure, bees associate floral traits with specific rewards and exhibit strong flower constancy, even in the presence of alternative options.[32] This constancy reflects not cognitive limitation but an optimization of memory recall, motor patterns, and search efficiency.[33] Bees also demonstrate spatiotemporal learning, remembering not only which flowers are profitable but also when and where rewards are available.[34] These memory-driven behaviors show that floral selection is not random but instead shaped by cognitive investment, highlighting the role of neural architecture in ecological interactions.

Pollination's Role in the Ecosystem:

Pollination significantly enhances both agricultural productivity and ecosystem stability. Klein *et al.* evaluated the world's leading crops and found pollinators to be essential or highly significant for many fruits, nuts, and oilseeds.[35] This dependence has increased in recent decades as agriculture has shifted toward pollinator-dependent crops.[36] At field scales, a landmark global analysis showed wild insects consistently increase fruit set, and honey bees do not substitute for wild pollinators; both groups act additively.[37] In the U.S., a 131-site, 7-crop study found that five of seven crops were pollinator-limited, with wild bees often matching honey bees in delivered pollination.[38] Landscape context is also critical. The synthesis across continents shows that both pollinator richness and visitation rates

decline exponentially as the distance from natural or semi-natural habitats increases, with visitation dropping by half within approximately 0.6 km, underscoring the direct role of habitat conservation in maintaining stable crop yields.[39] At broader scales, yield growth and stability are lower in pollinator-dependent crops, consistent with pollen limitation constraining production.[40]

Pollination supports plant growth and reproduction, thereby sustaining entire ecosystems. Because most flowering plants rely on animal pollination, declines in pollinator populations can propagate through plant–pollinator networks and affect wider food webs. Network theory and simulations by Memmott *et al.*[41] and Kaiser-Bunbury *et al.*[42] demonstrate that removing pollinators can trigger secondary plant extinctions and community collapse, with the extent of vulnerability depending on network structure. Empirical studies support these predictions. Biesmeijer *et al.*[43] documented simultaneous declines in bees, hoverflies, and their associated plant guilds, while Burkle, Marlin, and Knight, using a 120-year dataset, showed degraded network structure and the extirpation of ~50% of bee species, consistent with reduced pollination services for many plants.[44] Experimental manipulations provide mechanistic insight: removing a single pollinator species reduced the floral fidelity of remaining pollinators and depressed plant reproduction, causing cascading effects on seed and fruit resources for herbivores and frugivores.[45] Meta-analyses and reviews such as Knight *et al.*[46] further confirm that pollen limitation is widespread and can constrain plant demography, thereby affecting the quantity and timing of resources (seeds, fruits, foliage) that higher trophic levels depend on. Importantly, restoration experiments demonstrate that these disruptions are reversible: vegetation restoration strengthened network structure and improved pollination outcomes.[47]

Given that *Wolbachia* can influence fruit fly behavior, it is plausible that it may also affect *Anthophila* bees and their complex behaviors.[1,2] *Drosophila* and *Anthophila* share homologous brain structures, meaning they possess evolutionarily conserved regions responsible for similar functions.[12,14] Both groups also include species naturally infected with *Wolbachia* strains.[1,4,5,7] However, whereas fruit fly behaviors are largely limited to reproduction and social interactions, bees exhibit far more sophisticated behaviors. With the same brain structures, bees engage in activities beyond reproduction and communication, most notably, pollination.[15,18]

Pollination plays a critical role in ecosystems. The vast majority of flowering plants rely on bees for reproduction, which in turn supports plant propagation and ecosystem productivity.[35,36] Through pollination, plants reproduce and thrive, increasing the abundance of primary producers within ecosystems.[39,40] These producers form the foundation of the food chain, supplying energy to all consumer organisms. Because many species depend on plants for energy, and those plants often rely on bees for reproduction, bees play an essential role in maintaining ecosystem energy flow (Figure 4).[37,41,43]

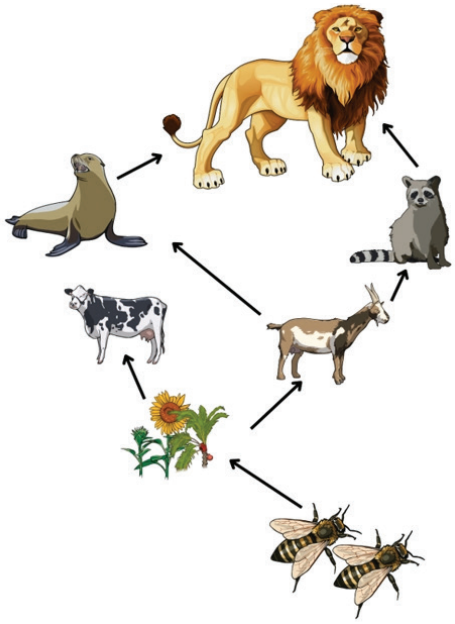


Figure 4: Flowers and plants, reproduced by pollinators, nourish the ecosystem with energy. The figure shows how this energy is transferred through the predation of species within the ecosystem. Initially, the flowers containing this energy can be consumed by primary consumers, such as cows or goats. Those animals are then consumed by secondary consumers, like raccoons or seals, and ultimately by tertiary consumers.

Therefore, it can be hypothesized that *Wolbachia* targets specific regions of the *Anthophila* bee brain that play key roles in pollination. If this is the case, pollination behavior could change, potentially changing ecosystem dynamics in a significant way. However, there is still little compelling scientific evidence to support this hypothesis. Very little is known about how *Wolbachia* spreads within *Anthophila* bees at the molecular level, or how it might influence behavior in real-world ecological settings.

Possible Effects of** Wolbachia **Infections on Pollination and Ecosystems:

Wolbachia may infect the brains of *Anthophila* bees by targeting such regions as antennal lobes, optic lobes, and mushroom bodies, structures critical for pollination. Such infections could lead to changes in pollination behavior. This hypothesis is supported by previous studies that examined the effects of *Wolbachia* on *Drosophila* fruit fly behavior, molecular evidence detecting *Wolbachia* in *Anthophila* bees, and the well-documented ecological importance of pollination.

As *Drosophila* and *Anthophila* share homologous brain structures, they possess evolutionarily conserved regions responsible for similar functions. Both groups also include species naturally infected with *Wolbachia* strains. Given that *Wolbachia* can influence fruit fly behavior, it can also affect *Anthophila* bees and their behaviors. However, while fruit fly behaviors are largely limited to reproduction and social interactions, bees exhibit much more sophisticated behaviors. With the same brain structures, bees engage in activities beyond reproduction and communication, most notably, pollination.

Pollination also plays a vital role in ecosystems. The vast majority of the plants in the ecosystem generally rely on bees

DOI: 10.36838/IJHSR88.8

for pollination. Not only are bees able to increase agricultural productivity, but they also support ecological networks by facilitating plant reproduction. Through pollination, plants can reproduce and thrive, increasing the abundance of primary producers within ecosystems. Producers form the foundation of the food chain, supplying energy to all consumer organisms. Since many species depend on plants for energy, and those plants often rely on bees for reproduction, the role of bees in maintaining energy flow is essential.

Currently, a direct causal role of *Wolbachia* in shaping pollination behavior has not yet been established, as the studies discussed here do not employ germ-free or antibiotic-mediated *Wolbachia* removal approaches in *Anthophila*. Pietri and colleagues reviewed *Wolbachia*'s somatic interactions and noted that antibiotic treatment of non-*Wolbachia*-infected *Drosophila* has substantial long-term effects on host physiology and behavior, including mitochondrial dysfunction and metabolic changes after tetracycline treatment. This demonstrates that antibiotics commonly used to clear *Wolbachia* (e.g., tetracycline) can independently impact host cell physiology beyond *Wolbachia* removal, complicating the interpretation of "germ-free."

Conclusion

The review of existing literature suggests that it is plausible to hypothesize that *Wolbachia* may infect specific structures of the *Anthophila* bee brain that are heavily involved in pollination. In this case, pollination patterns could be disrupted, significantly influencing ecosystem dynamics. Therefore, this study suggests the following two studies. First, molecular studies need to be conducted to explore the mechanisms by which *Wolbachia* localizes in the brains of *Anthophila* bees. It is important to identify the specific brain regions targeted by *Wolbachia* and determine the functions these regions control. In addition, the molecular mechanisms underlying these interactions require further investigation. For instance, *Wolbachia* proteins concentrated in neural tissues may alter transcription factors regulating brain-associated genes or disrupt signal transduction pathways critical for normal brain function.

Second, an analysis of the effects of *Wolbachia* infection on the pollen-collection behavior of *Anthophila* bees is needed. It is important to identify the types of flowers *Wolbachia*-infected bees prefer to pollinate, and the duration of time they spend probing each flower. The observations could provide useful insights into how *Wolbachia* infection influences pollination dynamics and impacts the overall ecosystem flow through agricultural productivity and changes to energy access in the food web.

Acknowledgments

I would like to thank my mentor, Dr. Grant Hartzog of UCSC, for his invaluable guidance throughout my research.

References

1. Porter, J., & Sullivan, W. (2023). The cellular lives of *Wolbachia*. Nature Reviews Microbiology, 21, 750–766. https://doi.org/10.1038/s41579-023-00918-x
2. Wang, J., & Wang, Y.-F. (2019). The effect of the endosymbiont *Wolbachia* on the behavior of insect hosts. Insect Science, 27(6), 846–858. https://doi.org/10.1111/1744-7917.12731
3. Strausfeld, N. J. (2012). Arthropod brains: Evolution, functional elegance, and historical significance. Cambridge/London: Belknap Press of Harvard University Press. https://doi.org/10.2307/j.ctv1dp0v2h
4. Stouthamer, R., Breeuwer, J. A. J., & Hurst, G. D. D. (1999). *Wolbachia* Pipientis: Microbial Manipulator of Arthropod Reproduction, 53. 71–102. https://doi-org-ssl.oca.korea.ac.kr/10.1146/annurev.micro.53.1.71
5. Pietri, J. E. DeBruhl, H., & Sullivan, W. The rich somatic life of *Wolbachia*. MicrobiologyOpen, 5(6), 923–936. https://doi.org/10.1002/mbo3.390
6. Tram, U., Ferree, P. M., & Sullivan, W. (2003). Identification of *Wolbachia*–host interacting factors through cytological analysis, Microbes and Infection, 5(11), 999–1011. https://doi.org/10.1016/S1286-4579(03)00192-8
7. Albertson, R., Tan, V., Leads, R., Reyes, M., Sullivan, W., & Casper-Lindley, C. (2013). Mapping *Wolbachia* distributions in the adult *Drosophila* brain. Cellular Microbiology, 15(9), 1527–1544. https://doi.org/10.1111/cmi.12136
8. Chen, M.-Y., Li, D., Wang, Z.-N., Xu, F.-Z., Feng, Y.-W., Yu, Q.-L., Wang, Y.-Y., Zhang, S., & Wang, Y.-F. (2024). Infection by virulent wMelPop *Wolbachia* improves learning and memory capacity in *Drosophila melanogaster*. Animal Behaviour, 212, 101–112. https://doi.org/10.1016/j.anbehav.2024.03.016
9. Warecki, B., Vega, G., Fowler, S., Hartzog, G., Karr, T. L., & Sullivan, W. (2025). *Wolbachia*-mediated reduction in the glutamate receptor mGluR promotes female promiscuity and bacterial spread. Cell Reports, 44(5), 115629. https://doi.org/10.1016/j.celrep.2025.115629
10. Gerth, M., Röthe, J., & Bleidorn, C. (2013). Tracing horizontal *Wolbachia* movements among bees (*Anthophila*): A combined approach using multilocus sequence typing data and host phylogeny. Molecular Ecology, 22(24), 6149–6162. https://doi.org/10.1111/mec.12549
11. Giurfa, M. (2007). Behavioral and neural analysis of associative learning in the honeybee: A taste from the magic well. Journal of Comparative Physiology A, 193(8), 801–824. https://doi.org/10.1007/s00359-007-0235-9
12. Strausfeld, N. J., & Hirth, F. (2013). Deep homology of arthropod central complex and vertebrate basal ganglia. Science, 340(6129), 157–161. https://doi.org/10.1126/science.1231828
13. Strausfeld, N. J., & Reisenman, C. E. (2009). Dimorphic olfactory lobes in the Arthropoda. Annals of the New York Academy of Sciences, 1170, 487–496. https://doi.org/10.1111/j.1749-6632.2009.04020.x*Drosophila*
14. Pfeiffer, K., & Homberg, U. (2014). Organization and functional roles of the central complex in the insect brain. Annual Review of Entomology, 59, 165–184. https://doi.org/10.1146/annurev-ento-011613-162031
15. Menzel, R., & Giurfa, M. (2001). Cognitive architecture of a mini-brain: The honeybee. Trends in Cognitive Sciences, 5(2), 62–71. https://doi.org/10.1016/S1364-6613(00)01601-6
16. Robertson, H. M., & Wanner, K. W. (2006). The chemoreceptor superfamily in the honey bee, *Apis mellifera*: Expansion of the odorant, but not gustatory, receptor family. Genome Research, 16(11), 1395–1403. https://doi.org/10.1101/gr.5057506

17. Berry, J. A., Cervantes-Sandoval, I., Nicholas, E. P., & Davis, R. L. (2012). Dopamine is required for learning and forgetting in *Drosophila*. Neuron, 74(3), 530–542. https://doi.org/10.1016/j.neuron.2012.04.007

18. Chittka, L., & Menzel, R. (1992). The evolutionary adaptation of flower colours and the insect pollinators' colour vision. Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 171(2), 171–181. https://doi.org/10.1007/BF00188925
19. Whitney, H. M., Chittka, L., Bruce, T. J. A., & Glover, B. J. (2009). Conical epidermal cells allow bees to grip flowers and increase foraging efficiency. Current Biology, 19(11), 948–953. https://doi.org/10.1016/j.cub.2009.04.051
20. Giurfa, M., Núñez, J., Chittka, L. *et al.* (1995). Colour preferences of flower-naive honeybees. J Comp Physiol A 177, 247–259. https://doi.org/10.1007/BF00192415
21. Lehrer, M., Horridge, G. A., Zhang, S. W., & Gadagkar, R. (1995). Shape vision in bees: Innate preference for flower-like patterns. Philosophical Transactions of the Royal Society B: Biological Sciences, 347(1320), 123–137. https://doi.org/10.1098/rstb.1995.0017
22. Galizia, C. G., & Rössler, W. (2010). Parallel olfactory systems in insects: Anatomy and function. Annual Review of Entomology, 55, 399–420. https://doi.org/10.1146/annurev-ento-112408-085442
23. Menzel, R. (2001). Searching for the memory trace in a mini-brain, the honeybee. Learning & Memory, 8(2), 53–62. https://doi.org/10.1101/lm.38801
24. Eltz, T., Sager, A., & Lunau, K. (2005). Juggling with volatiles: Exposure of perfumes by displaying male orchid bees. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 191(7), 575–581. https://doi.org/10.1007/s00359-005-0603-2
25. Giurfa, M., & Sandoz, J.-C. (2012). Invertebrate learning and memory: Fifty years of olfactory conditioning of the proboscis extension response in honeybees. Learning & Memory, 19(2), 54–66. https://doi.org/10.1101/lm.024711.111
26. Roulston, T. H., Cane, J. H., & Buchmann, S. L. (2000). What governs protein content of pollen: Pollinator preferences, pollen–pistil interactions, or phylogeny? Ecological Monographs, 70(4), 617–643. https://doi.org/10.1890/0012-9615(2000)070[0617:WGPCOP]2.0.CO;2
27. Waddington, K. D., & Holden, L. R. (1979). Optimal foraging: On flower selection by bees. The American Naturalist, 114(2), 179–196. https://doi.org/10.1086/283428
28. Greggers, U., & Menzel, R. (1993). Memory dynamics and foraging strategies of honeybees. Behavioral Ecology and Sociobiology, 32(1), 17–29. https://doi.org/10.1007/BF00172219
29. Harder, L. D. (1983). Functional differences of the proboscides of short- and long-tongued bees (*Hymenoptera*, Apoidea). Canadian Journal of Zoology, 61(7), 1580–1586. https://doi.org/10.1139/z83-212
30. Inouye, D. W. (1980). The terminology of floral larceny. Ecology, 61(5), 1251–1253. https://doi.org/10.2307/1936841
31. Laverty, T. M. (1994). Bumblebee learning and flower morphology. Animal Behaviour, 47(3), 531–545. https://doi.org/10.1006/anbe.1994.1077
32. Chittka, L., Thomson, J. D., & Waser, N. M. (1999). Flower constancy, insect psychology, and plant evolution. Naturwissenschaften, 86(8), 361–377. https://doi.org/10.1007/s001140050636
33. Gegear, R. J., & Laverty, T. M. (2001). Flower constancy in bumblebees: A test of the trait variability hypothesis. Animal Behaviour, 62(3), 689–696. https://doi.org/10.1006/anbe.2001.1804
34. Pahl, M., Zhu, H., Tautz, J., & Zhang, S. (2007). Circadian timed episodic-like memory—a bee knows what to do when, and also where. Journal of Experimental Biology, 210(20), 3559–3567. https://doi.org/10.1242/jeb.007492
35. Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences, 274(1608), 303–313. https://doi.org/10.1098/rspb.2006.3721
36. Aizen, M. A., Garibaldi, L. A., Cunningham, S. A., & Klein, A. M. (2009). How much does agriculture depend on pollinators? Lessons from long-term trends. Trends in Ecology & Evolution, 24(5), 203–209. https://doi.org/10.1016/j.tree.2008.12.007
37 Garibaldi, L. A., Steffan-Dewenter, I., Winfree, R., Aizen, M. A., Bommarco, R., Cunningham, S. A., Kremen, C., Carvalheiro, L. G., Harder, L. D., Afik, O., Bartomeus, I., Benjamin, F., Boreux, V., Cariveau, D., Chacoff, N. P., Dudenhöffer, J. H., Freitas, B. M., Ghazoul, J., Holzschuh, A., Klein, A. M., ... & Tscharntke, T. (2013). Wild pollinators enhance fruit set of crops regardless of honey bee abundance. Science, 339(6127), 1608–1611. https://doi.org/10.1126/science.1230200
38. Reilly, J. R., Artz, D. R., Biddinger, D., Bobiwash, K., Boyle, N. K., Brittain, C., ... & Winfree, R. (2020). Crop production in the USA is frequently limited by a lack of pollinators. Proceedings of the Royal Society B: Biological Sciences, 287(1931), 20200922. https://doi.org/10.1098/rspb.2020.0922
39. Ricketts, T. H., Regetz, J., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., Bogdanski, A., Gemmill-Herren, B., Greenleaf, S. S., Klein, A. M., Mayfield, M. M., Morandin, L. A., Ochieng, A., Potts, S. G., & Viana, B. F. (2008). Landscape effects on crop pollination services: Are there general patterns? Ecology Letters, 11(5), 499–515. https://doi.org/10.1111/j.1461-0248.2008.01157.x
40. Garibaldi, L. A., Aizen, M. A., Klein, A.-M., Cunningham, S. A., & Harder, L. D. (2011). Global growth and stability of agricultural yield decrease with pollinator dependence. Proceedings of the National Academy of Sciences, 108(14), 5909–5914. https://doi.org/10.1073/pnas.1012431108
41. Memmott, J., Waser, N. M., & Price, M. V. (2004). Tolerance of pollination networks to species extinctions. Proceedings of the Royal Society B: Biological Sciences, 271(1557), 2605–2611. https://doi.org/10.1098/rspb.2004.2909
42. Kaiser-Bunbury, C. N., Muff, S., Memmott, J., Müller, C. B., & Caflisch, A. (2010). The robustness of pollination networks to the loss of species and interactions: A quantitative approach incorporating pollinator behaviour. Ecology Letters, 13(4), 442–452. https://doi.org/10.1111/j.1461-0248.2009.01437.x
43. Biesmeijer, J. C., Roberts, S. P. M., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., Schaffers, A. P., Potts, S. G., Kleukers, R., Thomas, C. D., Settele, J., & Kunin, W. E. (2006). Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. Science, 313(5785), 351–354. https://doi.org/10.1126/science.1127863
44. Burkle, L. A., Marlin, J. C., & Knight, T. M. (2013). Plant–pollinator interactions over 120 years: Loss of species, co-occurrence, and function. Science, 339(6127), 1611–1615. https://doi.org/10.1126/science.1232728
45. Brosi, B. J., & Briggs, H. M. (2013). Single pollinator species losses reduce floral fidelity and plant reproductive function. Proceedings of the National Academy of Sciences, 110(32), 13044–13048. https://doi.org/10.1073/pnas.1307438110
46. Knight, T. M., Steets, J. A., Vamosi, J. C., Mazer, S. J., Burd, M., Campbell, D. R., Dudash, M. R., Johnston, M. O., Mitchell, R. J., & Ashman, T.-L. (2005). Pollen limitation of plant reproduction:

DOI: 10.36838/IJHSR88.8

Pattern and process. Annual Review of Ecology, Evolution, and Systematics, 36, 467–497. https://doi.org/10.1146/annurev.ecolsys.36.102403.115320
47. Kaiser-Bunbury, C. N., Mougal, J., Whittington, A. E., Valentin, T., Gabriel, R., Olesen, J. M., & Blüthgen, N. (2017). Ecosystem restoration strengthens pollination network resilience and function. Nature, 542(7640), 223–227. https://doi.org/10.1038/nature21071

Authors

Sean Park is passionate about saving lives. He is a certified Canadian national lifeguard with major interests in genetics and molecular biology. His research focuses on bacterial transmission and neuroscience.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Machine Learning-Assisted Airfoil Aerodynamics Performance Prediction

Jason Mak

West Windsor-Plainsboro High School South, 346 Clarksville Rd, Princeton, New Jersey, 08540, USA; jjson.mak@gmail.com

ABSTRACT: Airfoil design plays an important role in determining how efficiently an object travels through the air. It has numerous applications in the industry, ranging from the construction of commercial aircraft to racecars, where small improvements in lift and drag can significantly reduce fuel consumption and improve stability. This study applies machine learning principles to develop a computationally efficient approach for predicting the lift and drag performance of airfoils. Accurate prediction of lift and drag coefficients from computational fluid dynamics software requires computationally expensive solutions of the Navier-Stokes equations. To overcome this limitation, we developed a neural network-based alternative that significantly reduces computational time while maintaining prediction accuracy. First, we generate a training data set with labels using traditional computational fluid dynamics software (CFD). Then we train a convolutional neural network model to make predictions of the aerodynamic coefficients. The results demonstrate that neural networks can effectively learn the complex relationships between airfoil geometry and aerodynamic properties, providing rapid lift and drag predictions not requiring full numerical solution of the Navier-Stokes equations.

KEYWORDS: Physics and Astronomy, Theoretical, Computational and Quantum Physics, Fluid Dynamics, Neural Networks, Airfoils.

Introduction

Neural Networks and Deep Learning:

Machine learning and deep learning are gaining popularity in all disciplines of science and engineering.[1-5] Its cognitive and predictive capability lies in its ability to model and form non-linear responses to complex patterns. Humans refer to such skills as learning from experience. One such area is solving prognostic physics problems, where we set up mathematical equations to solve and predict the evolution of a system and then compute diagnostic information from the state of the system to help us evaluate the model results. This traditional approach is computationally expensive because numerically solving these mathematical equations often requires iteratively solving large matrices through time stepping until satisfactory solutions are found.

At the heart of machine learning is the non-linear neural network, which stores the weights learned during training. We can think of these weights as the experience learnt by the model. A neural network generally consists of three components: the input, the neurons, and the output. The neurons contain weights, biases, and activation functions that are used to transform the input into the output. The input can be thought of as a vector that describes the features of a problem, such as the pixel values of an image, the frequency and sound level values of an audio, etc. Based on the features that capture the main characteristics of the input, the neural network can calculate a set of desired outputs, e.g., the handwritten digit in the image or the mood expressed by the sound.

In the supervised learning approach, the neural network first undergoes a training process. Pre-selected input and output pairs are used to update the weights and biases stored in the neurons until the difference between the given output and the calculated output from the neural network is minimized. This process is similar to what occurs in a gradient descent algorithm. The quality of the training dataset (the input and output pairs) determines the quality of the trained neural network. After the neural network has been trained and verified, we can begin using it to make predictions for inputs that have not been seen by the neural network. Because the neural network calculations during this stage involve the forward multiplication of matrices and vectors, predictions can be made quickly. In the conventional approach, numerical solvers designed to make predictions involve iterative processes, such as the Navier-Stokes equation solver in computational fluid dynamics. This causes the conventional prediction approach to be less efficient and slower than the predictions performed by the trained neural network.[6-8]

A problem that is of interest is to apply such a supervised learning approach to make predictions of airfoil design in aerodynamics quickly.[9,10] Although neural network predictions may not be as precise as dedicated CFD software or results from wind tunnel measurements, they can provide quick feedback on whether a certain design is viable and suggest avenues for improving the design.[11] Such an approach is not limited to the airfoil problem; it has applications in many vision, audio, or signal-based systems, where we can rapidly improve the design by analyzing the image or audio signals.

DOI: 10.36838/IJHSR88.16

Methods

Convolutional Neural Network:

Traditionally, airfoil analysis is done by solving the Navier-Stokes equations in a meshed region around the airfoil with a prescribed free stream inlet and zero pressure gradient outlet boundary condition (Figure 1). The discrete mesh follows the geometry of the airfoil with a carefully generated boundary layer for flow analysis at the surface of the airfoil. The state variables pressure and velocity are calculated in each cell of the mesh during each time step until a quasi-equilibrium solution is obtained. These airfoil calculations are two-dimensional: assuming the wind is in the x-direction and height increases in the z-direction, from the pressure variables, the lift and drag coefficients can be computed by

$$C_l = \hat{z} \cdot \oiint P d\vec{A}$$

$$C_D = \hat{x} \cdot \oiint P d\vec{A}$$

Existing CFD software, such as Simflow and XFOIL, can calculate the lift and drag coefficients at different angles of attack for a given airfoil with reasonable speed on a personal computer.[12,13] Can we design a machine learning model to make predictions of the coefficients after it has been trained? It's a perfect problem where a complex relationship exists between the coefficients and the geometry of the airfoil and the angle of attack (AoA). The CFD models perform these predictions by directly solving the non-linear partial differential equations based on the geometry of the airfoil and the given angle of attack. If such a non-linear relationship exists from (airfoil, AoA) to lift and drag coefficients, the model should be able to learn it! The framework that allows us to learn from the geometry of an airfoil and angle of attack and predict the coefficients is called a convolutional neural network (CNN).[14]

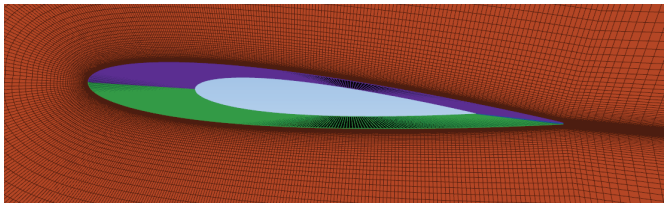


Figure 1: Generated mesh for airfoil NACA 0012 at an angle of attack of 5 degrees. Boundary layer cells are smaller than the cells further away. State variables pressure and velocity are computed in each cell in each time step. Diagnostic variables such as lift and drag coefficient are then computed from the pressure distribution. The green and purple parts of the mesh are on the frontal and rear surfaces of the airfoil.

We constructed a CNN model to analyze the airfoil images and trained the model using coefficients calculated from XFOIL. The convolutional layer can recognize the curvature and the tail of an airfoil through the various kernels and the pooling algorithm used. The convoluted data is then collapsed and concatenated with the angle of attack to form the input vector to the neural network. Figure 2 shows the CNN architecture used for machine learning. The input airfoil image is first convoluted with 32 kernels of size 3 and then max-pooled with a size of 2. Another iteration of convolution with 64 kernels of size 3 and maxpooling of size 2 is applied before the final arrays of images are collapsed with an angle of attack to form the input to the neural network. The neural network has an input layer, a dropout layer of 20%, another linear layer, and finally the output layer. To simplify analysis, we have chosen to focus on the lift coefficient only. Therefore, the output layer is of size 1.

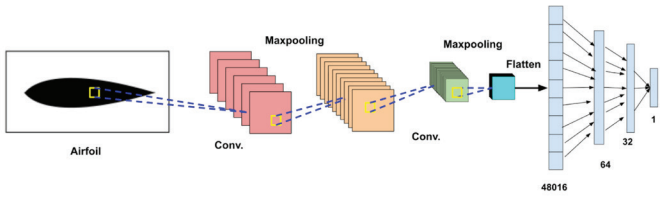
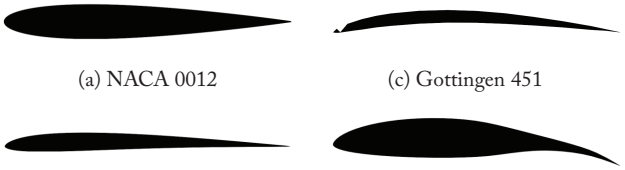


Figure 2: CNN architecture for the airfoil machine learning setup. The input airfoil image is converted into a monochrome black image. It's convoluted and pooled to create 64 channels of input and concatenated with the AoA to train the neural network. The airfoil image after the convolutional layer, plus 16 linear mapping of AoA, becomes 48016 input neurons to the neural network.

We downloaded the airfoil dataset from the official NACA database. The airfoil data are then converted to images of the same size without distortion based on the airfoil data. Figure 3 shows some of the airfoil images used to train the neural network.



(a) NACA 0012 (c) Gottingen 451

(b) Drela AG04 (d) Eppler E664

Figure 3: Airfoil images generated from the NACA airfoil database.[15] Occasionally, the image generated from the airfoil data can be defective, as shown in (c) for the airfoil Gottingen 451. This could be a result of polygon self-intersection when there is an error in the coordinates. These images are checked and eliminated from the training dataset. The airfoils with a round leading edge are usually subsonic as in (a), (b), and (d).

Setting up XFOIL and Simflow:

To verify the aerodynamics calculation and prepare training data for the CNN, we set up XFOIL and Simflow to run under similar conditions. XFOIL is designed specifically for such simulations and is faster to run. Simflow provides an airfoil mode for aerodynamics calculations and is used to verify the results calculated from XFOIL. Table 1 shows the lift coefficients calculated from XFOIL and Simflow for the airfoil ag04 at various angles of attack. The results from XFOIL and Simflow are in good agreement.

Table 1: Comparison of calculated lift coefficients between XFOIL and Simflow for the AG04 airfoil at various angles of attack. As the angle of attack increases, we can see a monotonic increase in the lift coefficients for both XFOIL and Simflow. This agrees with the physics of airfoils for AoAs less than the stall angle, as with a higher angle of attack, more air will be deflected downwards—creating a higher pressure difference/lift.

Angle of Attack	XFOIL	Simflow
0.0	0.1764	0.18
1.0	0.2916	0.29
2.0	0.4068	0.40
3.0	0.5219	0.50
4.0	0.6368	0.61
5.0	0.7514	0.72
6.0	0.8659	0.82

During training, two NACA airfoils, GOE 451 and Marsden, were deleted from the NN training sample due to abnormally high lift coefficients given by Simflow and XFOIL. In both XFOIL and Simflow, we have used an incompressible steady-state solver. XFOIL does not provide many options in fine-tuning the calculation because it is specifically designed for such calculations. For Simflow, we have chosen the SIMPLE solver with a Reynolds-averaged Navier-Stokes (RANS) turbulent scheme. Simflow also allows for fine-tuning of the generated mesh with boundary layer cell thickness and length. We used the default settings from Simflow for the airfoil calculation. The airfoil radius is left as the default value of 20 meters, and a Mach number of 0.3 or 100 m/s is used as the free stream air velocity. The default kinematic viscosity of air, 1.51 * 10-5 m^2/s, is used in the viscous calculation. This is equivalent to a Reynolds number of

$$\frac{uL}{\nu} = \frac{100\ m/s\ \cdot\ 20\ m}{1.51\cdot 10^{-5} m^2/s} = 1.32\cdot 10^8$$

In theory, the lift coefficient does not depend on the Mach number; it should be a function of geometry and angle of attack only. In practice, the calculated lift coefficients from XFOIL and Simflow show dependencies on the Reynolds number used.

Results and Discussion

We trained the CNN model with 173 images, each for 21 angles of attack (from -10.0 to +10.0 degrees with a 1.0 degree interval), after discarding many airfoils from UIUC due to numerical convergence issues at various angles of attack in the XFOIL calculations. Most of the images used to train CNN have a round leading edge, which usually indicates subsonic flight. The training data for each input configuration are the lifting coefficients calculated using the XFOIL program.[16] We also ran Simflow to validate the results calculated from XFOIL. A total record of 3633 entries in the format (airfoil image, angle of attack, lift coefficient) is used to train the CNN. After the CNN is trained, we use the model to make predictions of airfoils at various angles of attack. The test data set includes approximately 200 airfoils chosen randomly from the UIUC database with an SRS. A comparison of the two sets of coefficients is shown in Figure 4. The predicted coefficients are generally in agreement with the corresponding coefficients calculated from XFOIL, with a correlation coefficient squared of 0.997, indicating excellent agreement between the machine learning predictions and traditional CFD methods. To quantify the benefit of this approach, for an airfoil set at a specific angle, traditional CFD simulations like Simflow average several minutes in order to reach numerical convergence. In contrast, the trained CNN required under 30 seconds to predict the lift coefficients for the entire test data set at the given angles of attack, corresponding to millisecond-level inference per case.

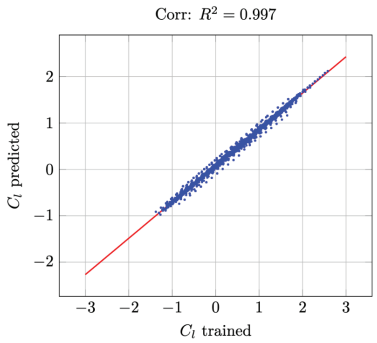


Figure 4: Comparison of predicted lift coefficients from XFOIL calculations and those obtained from the trained neural network across various airfoils from the UIUC database and different angles of attack. There is a clear linear trend with a best-fit line of y=0.783x+0.08. This trend is similar to that of the line y=x, representing perfect predictions of the lift coefficient.

Despite the high correlational coefficient, the CNN model exhibits a consistent systematic bias toward underestimating lift coefficients compared to XFOIL calculations. The regression analysis reveals a trend line slope of 0.783 with a y-intercept of 0.08. indicating that the model predictions are consistently lower than the reference values by approximately 22%. This underestimation trend is particularly evident across different airfoil geometries and angle of attack ranges, suggesting that the model may be conservative in its predictions. The systematic nature of this bias indicates that it could potentially be corrected through adjustment of the neural network architecture, though the underlying physical mechanisms causing this bias require further investigation.

The validation approach employed both XFOIL and Simflow calculations to ensure robustness of the reference data, with both tools configured for incompressible steady-state analysis at a Reynolds number around 1.32 * 10^8. While XFOIL and Simflow generally showed good agreement for most test cases, some discrepancies were observed, particularly for certain airfoil geometries. As shown in Table 1, the calculated results for the AG04 airfoil showed excellent agreement between XFOIL and Simflow across all angles of attack, with differences typically less than 5%. However, for other airfoils such as NACA 0015 and BW 050209, larger discrepancies were observed between the two CFD tools, highlighting the inherent uncertainties in numerical simulations and the importance of using multiple validation sources when training machine learning models for aerodynamic predictions.

DOI: 10.36838/IJHSR88.16

Table 2: Comparison of lift coefficients (Cl) from XFOIL, Simflow, and the trained neural network across various airfoils and different angles of attack. For this set of airfoils, the predicted values tend to be larger than the Simflow and XFOIL values; however, if we take into account the hundreds studied, there is a general trend of underestimation.

AoA	XFOIL	Simflow	Predicted	AoA	XFOIL	Simflow	Predicted
1.0	0.1163	0.223	0.2152	1.0	-0.0245	0.067	0.2132
2.0	0.2364	0.327	0.2936	2.0	0.0983	0.163	0.3000
3.0	0.3565	0.427	0.4137	3.0	0.2211	0.274	0.4288
4.0	0.4764	0.543	0.5051	4.0	0.3439	0.376	0.5216
5.0	0.5962	0.650	0.5934	5.0	0.4666	0.465	0.6098

(a) b737b (b) naca0015

AoA	XFOIL	Simflow	Predicted	AoA	XFOIL	Simflow	Predicted
1.0	0.5461	0.553	0.7424	1.0	0.2918	0.132	0.2876
2.0	0.6674	0.683	0.8527	2.0	0.4194	0.235	0.4256
3.0	0.7884	0.691	0.9606	3.0	0.5469	0.340	0.5212
4.0	0.9092	0.907	1.0651	4.0	0.6743	0.4407	0.6158
5.0	1.0297	0.924	1.1626	5.0	0.8014	0.538	0.7068

(c) goe701 (d) bw050209

AoA	XFOIL	Simflow	Predicted	AoA	XFOIL	Simflow	Predicted
1.0	0.3346	0.292	0.3359	1.0	0.6773	0.572	0.5897
2.0	0.4516	0.395	0.4036	2.0	0.795	0.671	0.6909
3.0	0.5685	0.496	0.4974	3.0	0.9126	0.768	0.7986
4.0	0.6852	0.594	0.5921	4.0	1.0298	0.861	0.9032
5.0	0.8017	0.689	0.6806	5.0	1.1467	0.946	1.0006

(e) rg14 (f) mh116

The model's performance varies significantly across different airfoil geometries, as illustrated in the comparative analysis (Table 2, Figure 5) of six representative airfoils. While some airfoils, like B737 B and RG 14, show relatively good agreement between predicted and calculated values, others exhibit more substantial deviations. The GOE 701 airfoil demonstrates the model's tendency to overestimate lift coefficients for high-lift airfoils. In contrast, symmetric airfoils like NACA 0015 show the model's difficulty in capturing the near-zero lift coefficients at low angles of attack. These airfoil-specific variations suggest that the current CNN architecture may benefit from additional geometric feature extraction layers to improve prediction accuracy across the full spectrum of airfoil designs.

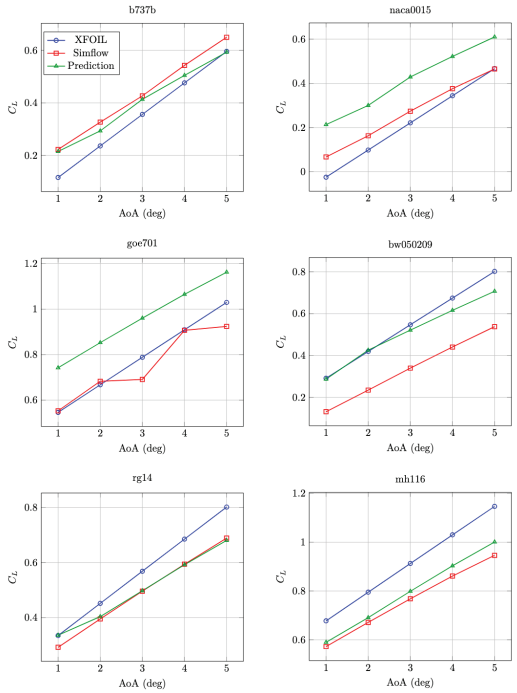


Figure 5: Comparison of lift coefficients obtained from XFOIL calculations (blue), Simflow calculations (red), and CNN predictions (green). The graphs indicate a linear trend except for some outliers (e.g., Simflow for GOE 701), suggesting a positive relationship between the lift coefficient and the angle of attack, which is physically true until the airfoil reaches the stall angle, where the lift starts to decrease.

Conclusion

In this work, we compare the lift coefficient predictions of a trained neural network against numerical calculations from XFOIL and Simflow using airfoils from the UIUC database over a range of angles of attack. The results demonstrated that the neural network can capture the overall trends in aerodynamic performance, often following the correct slope and relative magnitude of the lift curves. However, in several cases, the predicted values underestimate the reference data, as demonstrated by the scatterplot comparison in Figure 4. This underestimation highlights both the promise of the model and the need for further refinement to improve quantitative accuracy.

A key aspect of improving the model involves error analysis and mitigation. One immediate adjustment is to significantly increase the number of training epochs, as the relatively high validation error suggests that the network has not yet converged to an optimal solution. This could directly reduce the systematic underestimation observed in Figure 4. Additionally, adjusting the CNN model by adding more convolutional layers and changing the neural network layout could help improve the results.

Finally, the robustness of the results depends not only on model training but also on data integrity. Since visual inspection of CFD images or intermediate representations for errors is challenging, an automated verification mechanism may be required. One possible solution would be to train an auxiliary CNN tasked specifically with detecting corrupted or misaligned input images, thereby ensuring higher data quality for the main prediction network. While the present results are encouraging in demonstrating the feasibility of aerodynamic prediction with machine learning, careful attention to training strategies, model setup, validation practices, and data quality control will be essential for pushing the model toward application-ready performance.

Acknowledgments

I would like to express my gratitude to my mentor, Dr Fei Liu, for guiding me through this research and answering the many questions I had.

References

1. McCulloch, W. S.; Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133.
2. Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychol. Rev.* **1958**, *65*, 386–408.
3. Minsky, M.; Papert, S. *Perceptrons: An Introduction to Computational Geometry*; MIT Press: Cambridge, MA, **1969**.
4. Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533–536.
5. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
6. Kochkov, D.; Smith, J.; Alieva, A.; Wang, Q.; Brenner, M. P.; Hoyer, S. Machine Learning–Accelerated Computational Fluid Dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, e2101784118.

7. Vinuesa, R.; Brunton, S. L. Enhancing Computational Fluid Dynamics with Machine Learning. *Nat. Comput. Sci.* **2022**, *2*, 358–366.
8. Sharma, P.; Chung, W. T.; Akoush, B.; Ihme, M. A Review of Physics-Informed Machine Learning in Fluid Mechanics. *Energies* **2023**, *16*(5), 2343.
9. Li, J.; Du, X.; Martins, J. R. R. A. Machine Learning in Aerodynamic Shape Optimization. *Prog. Aerosp. Sci.* **2022**, *134*, 100849.
10. Loh, P. D.; Ong, O. A. Reinforcement Learning Approach to Airfoil Shape Optimization. *Sci. Rep.* **2023**, *13*, 9753.
11. Deters, R. W.; Williamson, G. A.; Selig, M. S. Wind Tunnel Testing Airfoils at Low Reynolds Numbers. In *49th AIAA Aerospace Sciences Meeting*; American Institute of Aeronautics and Astronautics: Orlando, FL, **2011**; Paper AIAA 2011-875.
12. *SimFlow*, Ver. 5.0; SimFlow Technologies, 2025.
13. Drela, M. XFOIL: An Analysis and Design System for Low Reynolds Number Airfoils. In *Low Reynolds Number Aerodynamics*, Mueller, T. J., Ed.; Springer: Berlin, Heidelberg, **1989**; pp 1–12.
14. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* **1998**, *86*, 2278–2324.
15. *UIUC Airfoil Coordinates Database*; Selig, M. S.; University of Illinois at Urbana-Champaign. https://m-selig.ae.illinois.edu/ads/coord_database.html
16. Drela, M.; Youngren, H. XFOIL, Ver. 6.99; Massachusetts Institute of Technology, 2013; http://web.mit.edu/drela/Public/web/xfoil/

Author

Jason Mak is a rising senior at West Windsor-Plainsboro High School South who enjoys learning physics and math. He hopes to pursue physics, engineering, or astronomy in college. In his free time, he plays badminton, tries new foods, stargazes, and reads the occasional book.

DOI: 10.36838/IJHSR88.16

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

The Prospect Theory Applied to Adolescents from Brazil

Laura Stumpf

Colegio Israelita Brasileiro, Av. Protasio Alves, 943, Porto Alegre, 90410-000, Brazil; laurastumpf15@icloud.com

ABSTRACT: Kahneman and Tversky developed their prospect theory in 1979, showing how the decision-making process in humans can be altered by biases called heuristics. Their theory changed forever how we think about behavioral economics and the expected utility theory. This research paper is intended to replicate the work of Kahneman and Tversky in the adolescent population. 13 adapted questions from the original work of Kahneman and Tversky were sent to high school adolescents from Porto Alegre, Brazil. 57 teenagers replied. The results showed statistical differences in 8 of the 13 questions. Although the adolescents usually answered in accordance with the work of Kahneman and Tversky, they were more prone to taking chances and risks to secure higher gains or avoid a sure loss. These results are significant because they demonstrate that teenagers are more susceptible to online gambling addiction and marketing manipulation. Further studies with larger populations are recommended.

KEYWORDS: Social and Behavioral Science, Behavioral Economics, Prospect Theory, Expected Utility.

Introduction

Daniel Kahneman and Amos Tversky[1] were pioneers in the science of behavioral economics. Their work showed how heuristics can affect the decision-making process, especially when those decisions may involve risk, challenging the Expected Utility Theory, which until then had been dominant in the academic field. From their studies, the "Prospect Theory" was developed, which explains how we make decisions differently when a situation has the potential to generate a loss or a gain. The research conducted by Kahneman and Tversky was carried out with university students and staff. Therefore, this paper aims to observe whether adolescents decide in the same way as adults when exposed to some of the same questions posed by Kahneman and Tversky in their 1979 work.

Before Prospect Theory, Expected Utility Theory was the model most widely used to explain decision-making in behavioral economics. The theory states that when people are confronted with situations of uncertainty, they choose the option that maximizes their expected utility, using the mathematical "expectation" operator. The utility of outcomes derives from their mathematical probability.

The pioneering work of Tversky and Kahneman demonstrated, in adults, that heuristics such as loss aversion (where the "disutility" of losing is greater than the utility of gaining) and the disposition effect (which says that investors sell winning assets too early and hold losing assets too long) influence decision-making when risk or gain is involved.

This paper aims to replicate part of the work of Kahneman and Tversky to determine whether adolescents in Brazil decide in the same way as adults when faced with uncertain situations.

1. Rational Expectations, Expected Utility, and the Prospect Theory:

1.1. Rational Expectations:

In classical economics, there is the assumption of rational expectations. The precursor of this theory was John Muth[2] in 1961, and Lucas[3] later refined it. The theory of rational expectations is based on the idea that people do not make systematic errors when predicting the future and make efficient use of all available information.[4] According to von Neumann and Morgenstern,[5] people are completely rational, use all available information efficiently, and markets are efficient. Thus, all decisions maximize expected utility.

By assuming that all economic agents have access to all available information, the theory argues that past errors do not influence present expectations. Therefore, systematic errors do not occur, and agents, on average, make correct predictions.[6]

Another important premise in economic and financial sciences is the liquidity preference theory, which states that economic agents and investors prefer liquidity and dislike risk. This theory was developed by the English economist John Maynard Keynes,[7] an important economic thinker. The theory states that risk-averse individuals prefer the certainty of a result over a risky outcome. Investors expect a premium to exchange liquid assets for long-term assets. This premium is the interest rate.[8]

However, not all scholars agree with the theory of rational expectations. According to authors such as psychologists Kahneman and Tversky,[1] human beings are often not rational in their decision-making. According to Daniel Kahneman's book Thinking, Fast and Slow,[9] the human brain works with two systems: one fast, intuitive, low-energy system that makes automatic decisions; and another that is analytical, slower, and where we allocate attention to complex cognitive tasks. According to the author, the fast system is the most used one to make our daily life more efficient. The result of this way of thinking is that our decisions are often affected by heuristics, intuitive solutions that can lead to mistakes. According to Tversky and Kahneman,[10] heuristics save time and energy and are generally effective, but they lead to systematic and predictable errors. Thus, humans would not make rational decisions, and a better

DOI: 10.36838/IJHSR88.21

understanding of heuristics and thought biases could improve decision-making in uncertain situations.

1.2. Expected Utility:

Until the advent of Prospect Theory, published in 1979 by Tversky and Kahneman,[1] the prevailing economic theory on how economic agents decide was Expected Utility Theory. The beginning of studies on expected utility dates back to the 18th century. At the time, decisions made under risk were seen as a rule that associated the values of available alternatives with their mathematical probabilities of occurrence, so that the decision-maker would choose the alternative with the greatest expected utility.[11] Utility was an indicator of the individual's satisfaction with the chosen option.[12]

Later, Daniel Bernoulli refined this theory to demonstrate the logic of human choices by formulating Decision Theory (1738). Bernoulli established a relationship between the psychological desirability of money and the actual amount of money.[9] He observed that most people dislike risk and aim to avoid unfavorable outcomes. The decision theory proposed by Bernoulli was still based on the principle of expected utility (the mathematical weighted average of outcomes). Still, he replaced the objective value scale with a subjective utility scale. He stated that people's choices are not based simply on the monetary value of bets, but on the psychological utility of outcomes, on their utility. Thus, Bernoulli refined expected utility theory, making it more plausible and possibly eliminating its errors.[11]

Utility is therefore not the weighted average of monetary effects, but the average of the utilities of those effects.[9] From Bernoulli's theory emerges the law of diminishing marginal utility, an important axiom of economic theory, which says that as wealth increases, the additional utility of one more unit of money becomes smaller, due to the increase in wealth itself. It would be something like saying: the utility of gaining USD$100 is greater when someone earns USD$2,000 per month than when someone earns USD$20,000 per month. In mathematical terms, this law says that utility as a function of money or wealth is a concave function (Figure 1).

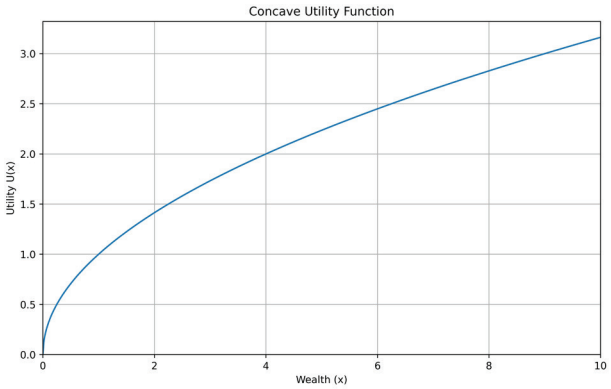


Figure 1: Graph exemplifying a concave utility function.

1.3. Reference Point:

However, according to Kahneman,[9] Bernoulli's model presents an important flaw: it lacks the idea of a reference point. Expected Utility Theory could explain risk aversion, but cannot explain the preference for risk-taking in betting — a behavior often observed in entrepreneurs and generals when all their options are bad.

In his book, Kahneman[9] gives an example of the problem of lacking a reference point: Anthony has 1 million dollars, and Betty has 4 million. Both are offered either a gamble or a sure gain: a 50% chance of ending up with a fortune of 1 million and a 50% chance of ending up with a fortune of 4 million, or a 100% certainty of having a fortune of 2 million dollars. Certainly, they will think differently, because for Betty, neither option is advantageous. She will certainly choose to gamble rather than lose half of her fortune for sure. Anthony, on the other hand, will probably choose to double his fortune with certainty rather than gamble and have a 50% chance of remaining with the same current fortune.

Expected Utility Theory assumes that decision-makers maximize their choices.[11] Despite its many inconsistencies, the theory was used by one of the greatest mathematicians of the 20th century, John Von Neumann, and economist Oskar Morgenstern,[5] who transformed the theory into an axiomatic model prescribing how rational economic agents should make their decisions. The authors used Expected Utility Theory in two ways: as logic determines decision-making and as a description of how economic agents decide.

1.4. Flaws in the Rational Expectations Theory

Authors such as Barreiros, Protil, and Moreira[13] argue that rational behavior in decision-making results in several difficulties, since other human perspectives within organizations began to be seen as political, social, and environmental behaviors. These other dimensions of human behavior and thought make it difficult for pure rationality to be achieved. According to Klacynski,[14] adults often misinterpret statistical base rates, overemphasize the value of vivid, readily available memories, make judgments based on misconstrued probability data, and violate the axioms of inference.

1.5. The Prospect Theory and Heuristics:

As psychologists and scholars of human behavior and heuristics, Kahneman and Tversky[1] questioned the absolute validity of this theory through studies conducted to elucidate how people decide when facing uncertainties. According to Silva and collaborators,[15] Kahneman and Tversky support the thesis of the presence of irrationality through the certainty effect, reflection effect, and isolation effect in decision-making to research the cognitive processes employed in non-rational decisions by individuals. Their work resulted in the Prospect Theory.[1]

■ Methods

A Google Form was sent to high school students at three private schools in Porto Alegre. It contained 13 questions adapted from Kahneman and Tversky's[1] work and an informed consent form. Not all questions of the original appear in Kahneman and Tversky's[1] were used to facilitate the application of the questionnaire, and some questions were presented in a different order to facilitate understanding. The payoffs

on the original paper were directly converted into Brazilian Reais. The consent form was sent through Google Classroom and explained that participation in the research was optional. Anonymity was guaranteed since the form with the questionnaire was stored separately from the consent form, and the filled questionnaire did not have any identifier. The research methodology was approved by the researcher's school ethics committee.

Each problem consisted of two options (A and B). In each option, there was a question involving the possibility of a gain or a loss. The students were instructed to choose between one of the two options (A or B).

The problems are described below:

PROBLEM 1: Choose between
- A: 33% chance of winning R$2,500 + 66% chance of winning R$2,400 + 1% chance of winning R$0
- B: 100% chance of winning R$2,400

PROBLEM 2: Choose between
- A: 80% chance of winning R$4,000
- B: certain gain of R$3,000

PROBLEM 3: Choose between
- A: 80% chance of losing R$4,000
- B: certain loss of R$3,000

PROBLEM 4: Choose between
- A: 50% chance of winning a 3-week trip through England, France, and Italy
- B: certain gain of a 1-week trip to England

PROBLEM 5: Choose between
- A: 5% chance of winning a 3-week trip through England, France, and Italy
- B: 10% chance of winning a 1-week trip to England

PROBLEM 6: Choose between
- A: 90% chance of winning R$3,000
- B: 45% chance of winning R$6,000

PROBLEM 7: Choose between
- A: 90% chance of losing R$3,000
- B: 45% chance of losing R$6,000

PROBLEM 8: Choose between
- A: 0.1% chance of winning R$6,000
- B: 0.2% chance of winning R$3,000

PROBLEM 9: Choose between
- A: 50% chance of winning R$1,000
- B: certain gain of R$500

PROBLEM 10: Choose between
- A: 50% chance of losing R$1,000
- B: certain loss of R$500

PROBLEM 11: Choose between
- A: 0.1% chance of winning R$5,000
- B: certain gain of R$5

PROBLEM 12: Choose between
- A: 0.1% chance of losing R$5,000
- B: certain loss of R$5

PROBLEM 13: Choose between
- A: 45% chance of winning R$6,000
- B: certain gain of R$3,00

After collecting the responses, the percentage of answers (A or B) for each problem was computed, the total percentage for each group was calculated, and then compared with the responses from Kahneman and Tversky's[1] original study. Each problem was analyzed individually, and the percentage of each response from each problem from this research was analyzed statistically, comparing to the reference sample (Kahneman and Tversky[1]). Data was analyzed using the chi-square test and Fisher's exact test. These tests compare two categorical distributions. For statistical analysis, the Jamovi Statistical software was used.

Results and Discussion

1. Results:

A total of 57 responses were received from the emails sent to the students.

Table 1 shows the results obtained from the students in the present study and the data collected in Kahneman and Tversky´s paper,[1] including the chi-square and Fisher´s Exact test p-values. A p-value lower than 0.05 was considered statistically significant. The results are also presented by each individual question, comparing the percentage of adolescents who chose options A or B in this study to the results from Kahneman and Tversky in Figures 2 to 14.

Table 1: Results (in percentage) from the adolescent population vs results from Kahneman and Tversky, and the statistical results of chi-square and Fisher tests.

	ADOLESCENTS		KAHNEMAN AND TVERSKY			
	A	B	A	B	CHI2 P-VALUE	FISHER P-VALUE
PROBLEM 1	42,1	57,9	18	82	0,00036	0,00019
PROBLEM 2	38,6	61,4	20	80	0,00625	0,00502
PROBLEM 3	70,2	29,8	92	8	0,00017	0,00010
PROBLEM 4	45,6	54,4	22	78	0.00072	0,00054
PROBLEM 5	67,7	32,3	67	33	1	1
PROBLEM 6	77,2	22,8	86	14	0,15462	0,14449
PROBLEM 7	29,8	70,2	8	92	0,00017	0,00010
PROBLEM 8	80,7	19,3	73	27	0,26134	0,2393
PROBLEM 9	52,6	47,4	16	84	0,00114	0,0049
PROBLEM 10	56,1	43,9	69	31	0,08211	0,0793
PROBLEM 11	78,9	21,1	72	28	0,3323	0,3239
PROBLEM 12	38,6	61,4	17	83	0,0011	0,0008
PROBLEM 13	21,1	78,9	14	86	0,2568	0,1971

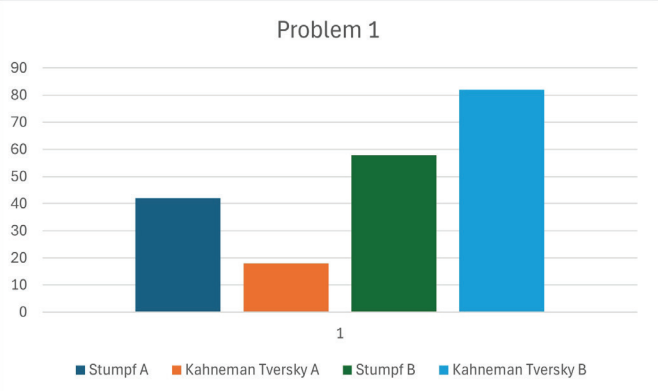


Figure 2: Results from problem 1 show a significant difference between the responses in this study and those in Kahneman and Tversky.

DOI: 10.36838/IJHSR88.21

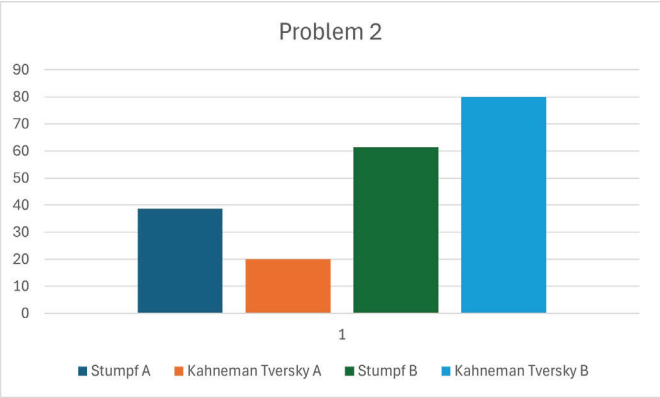


Figure 3: Results from problem 2. It shows that adolescents seek greater gain.

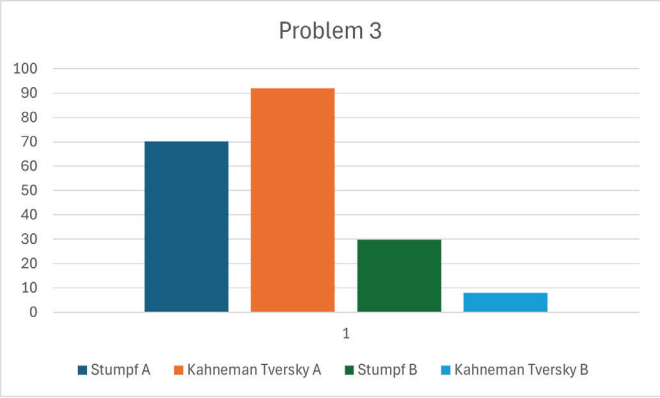


Figure 4: Results from problem 3. Here, there were significant differences because some adolescents prefer certain losses.

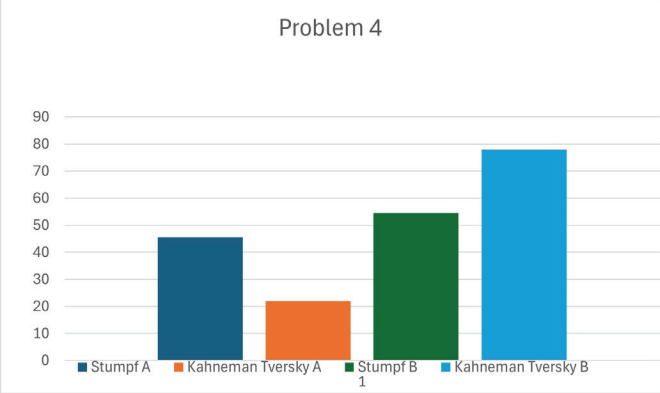


Figure 5: Results from problem 4 showed that teenagers try to secure better outcomes even when there's risk involved.

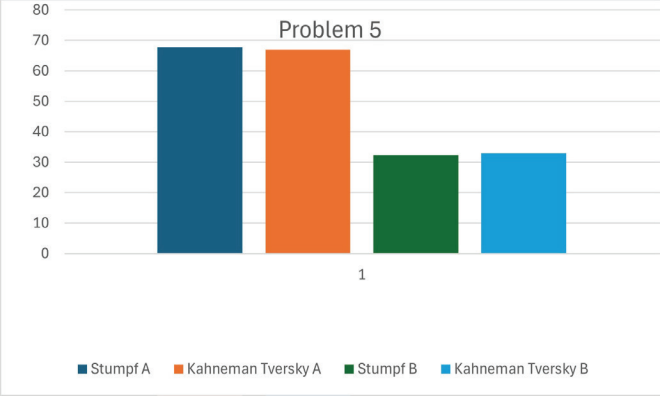


Figure 6: Results from problem 5. Here, adolescents behaved like adults, trying to obtain the best benefit.

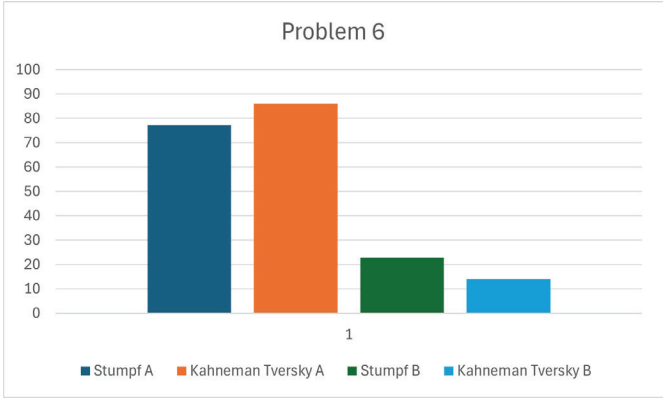


Figure 7: Results from problem 6 show that when the chance of winning is small, adolescents behave just like adults.

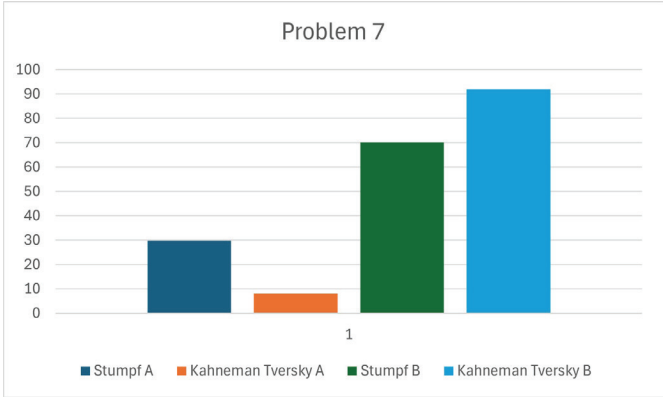


Figure 8: Results from problem 7 demonstrate that teenagers take less risk, trying to avoid greater losses.

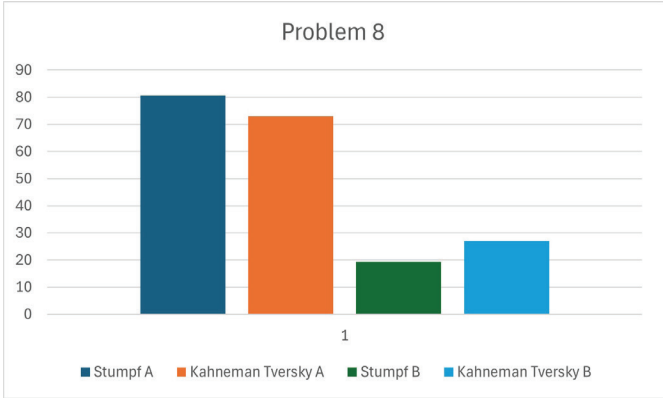


Figure 9: Results from problem 8 show significant differences between the two studies.

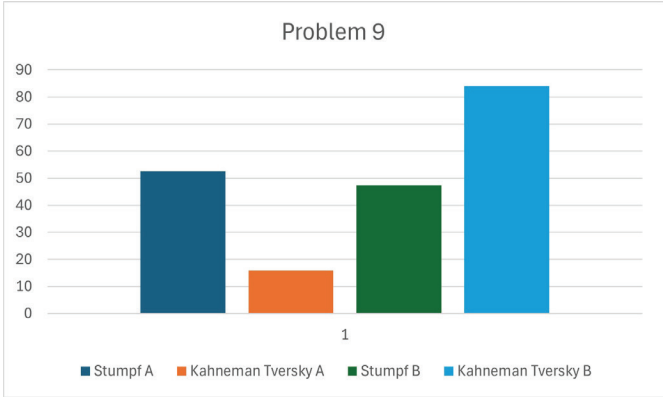


Figure 10: Results from problem 9. They show that teenagers take risks seeking greater rewards.

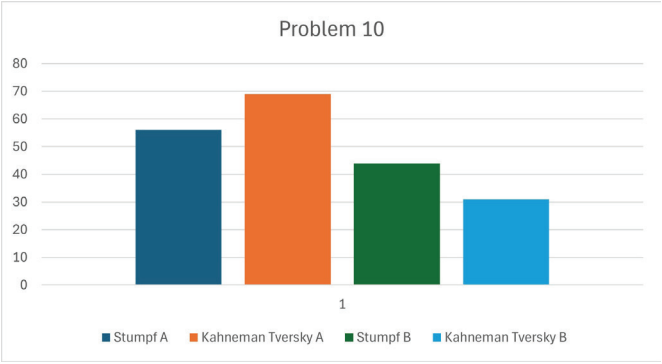


Figure 11: Results from problem 10 show there were no differences when adolescents and adults were trying to avoid losses.

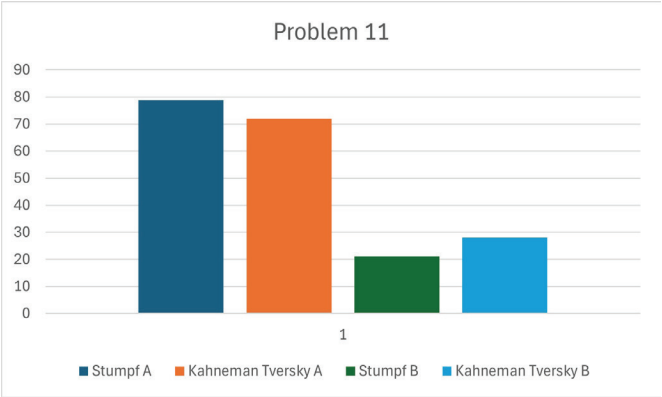


Figure 12: Results from problem 11. They show no significant difference in a problem simulating a lottery prize (a tiny chance of winning a lot).

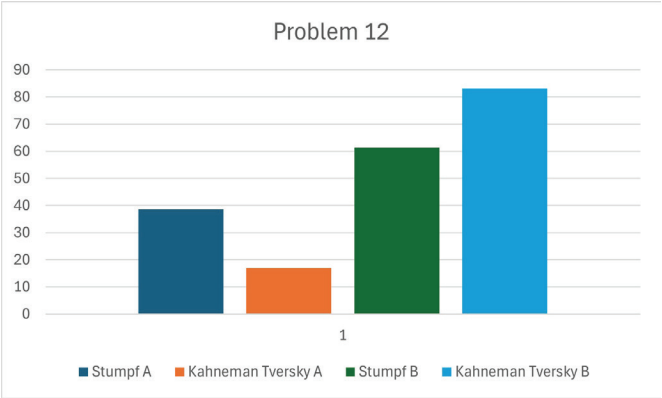


Figure 13: Results from problem 12. They demonstrate that adolescents try to avoid a sure loss.

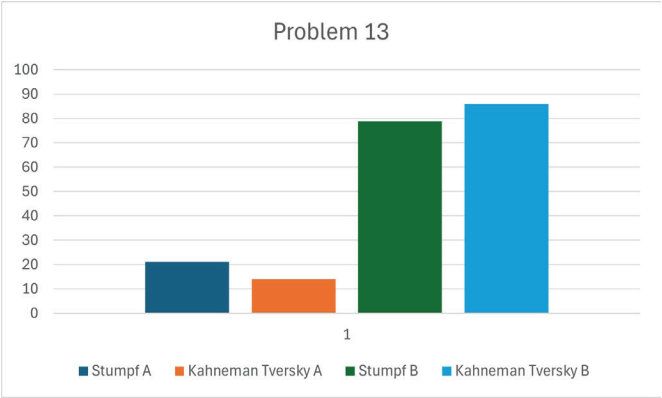


Figure 14: Results from problem 13. When the chances of winning are small, adolescents prefer a sure gain rather than risking, like adults.

2. Discussion:

Daniel Kahneman´s book "Thinking Fast and Slow"[9] explains that our reasoning process is formed by what he calls System 1 and System 2. System 1 is fast, intuitive, stores our memories, has low energy consumption, and tends to believe any information. System 2 is the system we use when doing calculations, analyses, or when we need to focus and exercise self-control. It is slower and consumes much more energy. To make daily life easier, human beings tend to rely primarily on System 1, which results in what we call heuristics.

Heuristics are simple and efficient mental strategies used in decision-making that, although not always leading to the best solution, work very well in everyday contexts. Some examples include:

• Representativeness heuristic: When we judge something only by how it appears, using insufficient information (since System 1 tends to believe any information), this is almost analogous to prejudice.

• Availability heuristic: When we evaluate the probability of an event based on how easily we can recall it.

• Affect heuristic: When emotions affect decisions (people in a good mood tend to decide one way, while those in a bad mood decide another). If something makes us feel good, we tend to judge it as low risk. The opposite occurs when something makes us feel bad.

• Anchoring heuristic: when irrelevant information influences the decision-making process.

• Disposition effect: investors tend to sell profitable assets too early and hold onto losing assets for too long, hoping to reverse the losses.

• Certainty effect: people are predisposed to give more weight to certain events than to probable ones, even when the difference in probability is small. Kahneman and Tversky based themselves on the study by Maurice Allais[16] to demonstrate the "certainty effect," proving that people give more importance to "certain" outcomes than to those involving uncertainty.

• Framing effect: when the way the information is presented affects people's decision-making process. The same information can lead to different choices depending on whether it is framed in terms of positive or negative outcomes.

2.1. Problem 1:

When analyzing problem 1, some differences between the samples can be observed. In Kahneman and Tversky's study[1] 82% of people chose option B, showing a violation of Expected Utility Theory. The population chose option B as if the utility of $U(2400) > 0.33U(2500) + 0.66U(2400)$, when in fact option A has a higher expected utility. In the present study, 57.9% of students chose option B, but 42.1% chose option A, a statistically significant result compared to Kahneman and Tversky's. This may indicate that some school-age students are more attentive to mathematical issues or more careful in calculating before responding, thus confirming Expected Utility Theory for 42.1% of the population.

DOI: 10.36838/IJHSR88.21

2.2. Problem 2:

In problem 2, subjects are faced with a certain gain or an 80% possibility of an even larger gain. In the Kahneman and Tversky study, 80% chose option B, even though the expected value of option A is identical. According to the authors, this demonstrates that people prefer the certainty of a gain; certainty is overestimated compared to uncertainty. In this study, 61.4% of students chose a certain gain, following the same trend but at a significantly lower rate. These results may indicate that younger people are more willing to take risks (considering the high 80% probability of winning) in search of a higher gain rather than settling for a certain gain.

The certainty effect defines the behavior where, in a bet involving positive outcomes, people prefer to avoid risk to secure a guaranteed gain. This was confirmed in this study, but at a significantly lower intensity among teenagers compared to adults, showing that a significant number of adolescents tend to seek risk trying to obtain higher gains, or maybe adolescents are more willing to seek the emotion of gambling than adults, or they did not fully understand the question.

2.3. Problem 3:

This case shows the reverse — the certainty of a loss. Here, people tend to take risks to try to avoid a sure loss. In Kahneman and Tversky's study, 92% chose option A, even though the probabilities were the same. They describe this as the reflection effect: when faced with gains, people are risk-averse; when faced with losses, they take risks to avoid a sure loss. Markowitz[17] had already suggested that decision-making focuses on losses or gains, not on absolute magnitudes.

In this study, the same behavior was confirmed, but with lower intensity: 70.2% chose to gamble to avoid a sure loss. This may indicate that teenagers either did not fully understand the question or that around 30% feared the possibility of an even greater loss compared to the certain one. Adolescents may even have a different interpretation of "chance" than adults.

Overall, both for positive and negative outcomes, Expected Utility Theory is violated. Adolescents, however, seemed more inclined to calculate the math involved, as in Question 1, where 42.1% chose option A (which presented the higher utility). These results also show that adolescents are more willing to take risks in search of larger gains, contradicting some of Kahneman and Tversky's findings[1].

2.4. Problems 4 and 5:

Problems 4 and 5 will illustrate the certainty effect of non-monetary outcomes. In the original study, 78% chose option B in Question 4, evidencing the preference for certain gains. In Question 5, 67% chose A in the original study, as there was no certain gain and people tended to select the more favorable outcome even with lower expected utility.

In this study, most adolescents also chose a certain gain in Question 4, but a significant portion opted to gamble for a more favorable outcome. Again, this implies adolescents' willingness to risk searching for better outcomes. In Question 5, their responses matched Kahneman and Tversky's sample, with no significant difference — once again suggesting that teenagers seek better outcomes and are more willing to gamble.

2.5. Problems 6, 7, and 8:

These show how Expected Utility Theory fails to capture behavior under risk fully.

In Question 6, 86% of Kahneman and Tversky's participants chose A, preferring the higher probability of winning. In this study, 77.2% chose A, which is not significantly different.

In Question 7, 92% in the original study chose B (taking the smaller probability of loss, even if the loss was larger). Here, 70.2% chose B — the same trend but at significantly lower rates. This may indicate that adolescents have noticed the mathematical equivalence or that they want to avoid a larger loss.

In Question 8, 73% in the original study chose A (when probabilities were very small, people picked the option with the larger prize). In the present study, results were statistically similar to theirs.

2.6. Problems 9 and 10:

Both problems 9 and 10 involve identical expected utilities.

In Question 9 (certain gain vs. possible higher gain), 84% in the original study chose the certain gain. Here, however, 52.6% of the students in this sample chose to gamble — contradicting the prospect theory, which says people avoid risks to secure gains but take risks to avoid losses. In the present study, adolescents preferred gambling, maybe trying to obtain a higher prize.

In Question 10 (certain loss vs. gambling to avoid loss), teenagers behaved like the original sample, but at a lower (not significant) intensity. This matched the prospect theory for losses, but not for gains.

Thus, in this sample, the prospect theory was confirmed only for losses. Teenagers may prefer to take risks to get higher gains.

2.7. Problems 11 and 12:

Both problems 11 and 12 involve a very small probability and are discussed below.

In Question 11, 72% in the original study chose the tiny chance of a big prize over a sure small gain — comparable to buying a lottery ticket. Adolescents answered similarly, with no statistical difference.

In Question 12, 83% of adults choose the sure small loss (like buying insurance). Teenagers responded similarly, but a statistically significant portion preferred to gamble instead of losing R$5 for sure. Possibly, they calculated the minimal probability of losing R$5,000 and preferred to take the risk, or maybe R$5.00 means more to an adolescent than to an adult, or they misunderstood the question.

2.8. Problem 13:

Again, problem 13 will address how Expected Utility Theory fails when people are faced with a certain gain. Mathematically, both options are identical, but both adults and teenagers choose a certain gain.

However, adolescents showed differences compared to earlier questions. In Question 2 (80% chance to win R$4,000 vs. a certain gain of R$3,000), many preferred to gamble; maybe the 80% probability of gain seemed high enough to justify gambling to adolescents. In Question 13 (45% chance to win R$6,000 vs. a certain R$3,000), most preferred the certain gain. The probability of gaining R$6,000 being below 50% may explain why these teenagers preferred the safer choice in this problem.

2.9. General Observations:

Kahneman[9] states that humans tend to focus on changes relative to a reference point. For example, losing 100 or 200 feels more significant than losing 1,100 or 1,200, even though the absolute loss is the same. Another point of Prospect Theory is that losses matter more than gains, and the value function is:

- depending on a reference point,
- concave for losses and convex for gains,
- steeper for losses than for gains.

Figure 15 illustrates the Prospect Theory curve, which is concave for losses and convex for gains, showing that we are risk-averse when securing a gain and risk seekers when trying to avoid losses. The steepness of the loss curve demonstrates that losses are more important to us than gains.

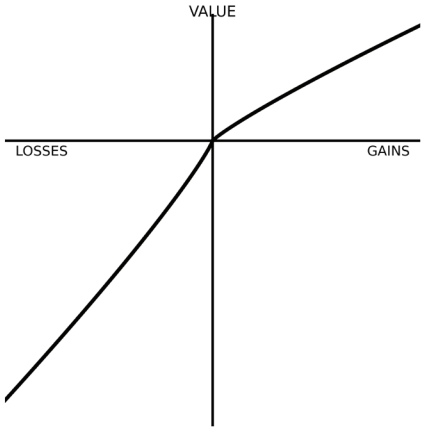


Figure 15: Prospect theory graph taken from Kahneman and Tversky´s paper.

In this study, reference points were not directly tested, but it can be inferred that the monetary amounts in the questions may seem larger for teenagers than for adults.

The results of this study show that adolescent behavior regarding gains and losses generally follows the trends of Prospect Theory. Still, in many of the proposed problems, the results were significantly different, with adolescents showing a higher disposition to gamble.

Chien *et al.*[18] tested 92 adolescents using Tversky and Kahneman decision problems for framing effects. The response pattern obtained by the authors resembled the general pattern exhibited by adults tested by Tversky and Kahneman. Negative frames led them to accept risk to avoid certain loss; positive frames prevented them from risking securing a certain gain. Chien *et al.*[18] only found the framing effect in some of the young adolescents, but the influence of the framing effect was consistent. This corroborates the findings in this study. A study by Rosebaum *et al.*[19] showed that adolescents are more prone to risk-taking than adults. The authors suggest that these point to a complex interaction between multiple psychological constructs that develop across adolescence. In this study, more than half of the questions had responses that differed significantly from those of adults, showing greater risk propensity. This propensity might be concerning, given teenagers' exposure to online gambling sites and online marketing. The tendency towards gambling obtained in this study can also be caused by a lack of proper understanding of the nature of the questions.

A 2024 Guardian[20] report stated that online betting sites are a public health problem, rapidly expanding via mobile phones and harming more people than previously thought. Teenagers are a vulnerable group, exposed to sophisticated marketing and gambling integrated into video games. Montiel *et al.*[21] reviewed studies on online gambling among youth aged 11–21 from the past 20 years and found that between 0.77% and 57.5% showed some pathological degree of gambling, depending on measurement instruments, sample, cut-off points, and criteria applied, and between 0.89% and 1% had a full disorder. They stated that, although illegal among adolescents, online gambling is a common practice and it´s expanding alongside the expansion of technology. 5-15% of adolescents gamble online, and 40-70% gamble offline, which puts their mental health and well-being at a serious risk.

Together with the present findings, which suggest teenagers are more risk-prone, these studies reinforce the need for regulation of online gambling for minors. In a meta-analysis by Ding *et al.*,[22] 40 relevant studies across 12 databases were reviewed, and the authors found that sensation seeking and risk taking are typical in adolescents, and the incidence of risky behaviors is significantly higher in teenagers. Adolescents' still-developing brain functions may lead to cognitive dissonance in decision-making. According to these authors, adults are more developed physically and mentally, and are more cautious when making risky decisions. Although the present study did not investigate online gambling among adolescents, but instead the choices they make when confronted with a gamble, a possible tendency towards risk was observed which may come to the encounter of the results found by Ding *et al.*[22] This might suggest that lawmakers should consider the adolescent decision-making process in matters such as marketing regulation and online gambling site regulation to protect adolescents' mental and financial well-being.

■ Limitations

This study shows some important limitations. Firstly, the obtained sample is relatively small (only 57 subjects participated in this study), especially when compared to Kahneman and Tverski[1] original work that used a larger sample. Furthermore, the sample was drawn entirely from three private schools in a single city, which may not represent adolescents in general. Students from private schools usually come from affluent families; hence, their relationship with monetary values may differ from that of students from public schools, especially in Brazil, where socioeconomic inequality is profound. Both limitations

DOI: 10.36838/IJHSR88.21

may jeopardize the possibility of generalization of these results to broader populations. Further studies with larger samples and samples drawn from different socioeconomic backgrounds are recommended.

Conclusion

The present study replicated some of the questions originally asked of adults in Kahneman and Tversky's work and applied them to high school adolescents in Brazil. Significant differences were found in more than half of the proposed problems.

It was shown that teenagers tend to take more risks to obtain more favorable outcomes, both in the face of gains and losses, contradicting some classic heuristics of Behavioral Economics Theory. This behavior may indicate that teenagers are more audacious and less responsible with money than adults. It may also reflect a greater search for excitement compared to adults.

These results might highlight the dangers of exposing teenagers to online gambling sites and manipulative marketing. We recommend that new studies with larger samples be conducted to confirm the trends identified in this study.

Acknowledgments

I thank Dr. Carlos P. Bergmann of Federal University of Rio Grande do Sul for helping me with the statistical analysis.

References

1. Kahneman, D. P.; Tversky, A. Prospect theory: an analysis of decision under risk Econometrica, 1979, 47, 263-292.
2. Muth, John, Rational expectations and the theory of price movements. Econometrica, 1961, 29, 315-335.
3. Lucas J.R., R. E. Expectations and the neutrality of money. Journal of Economic Theory, 1972, 4 (2), 103-104.
4. Santos, N.C.; A hipótese de expectativas racionais: a teoria e os testes. https://lume.ufrgs.br/bitstream/handle/10183/5475/000426935.pdf?sequence=1&isAllowed=y accessed Oct 13, 2024.
5. Von Neumann, J.; Morgenstern, O. Theory of games and economic behavior. 2a ed, Princeton University Press, 1947.
6. Curso Sapientia. Hipótese das expectativas racionais. https://www.cursosapientia.com.br/conteudo/noticias/dicionario-de-economia-para-o-cacd- hipótese-das-expectativas-adaptativas-e-racionais, accessed Sept, 17, 2024.
7. Keynes, J.M.; Teoria Geral do Emprego, do Juro e do Dinheiro, 1st ed., Abril Cultural, 1983.
8. Investopedia: Theory of liquidity preference: definition, history, how it works, and example. www.investopedia.com. Accessed Aug 31, 2025.
9. Kahneman, D. Thinking Fast and Slow. 1st ed, Farrar, Strauss and Giroux, 2013.
10. Tversky, A., Kahneman, D. Judgment under uncertainty - heuristics and biases, *Science*, **1974**, 185, 1124-1131. https://doi.org/10.1126/science.185.4157.1124.
11. Coombs, C. H.; Dawes, R. M.; Tversky, A. Mathematical Psychology: an Elementary Introduction. 1 ed, Prentice-Hall, 1970.
12. Varian, H. R. Microeconomia – princípios básicos. 7 ed, Elsevier, 2003.
13. Barreiros, R. F.; Protil, R. M.; Moreira, V. R. Caracterização da Natureza do Processo Decisório em Nível Estratégico nas Cooperativas Agroindustriais do Estado do Paraná. Em: Encontro da Associação de Pós-Graduação e Pesquisa em Administração, Brasília, Brasil, Set 17-21, 2005, Anais ANPAD, Atibaia, 2005.
14. Klaczynski, P.A.; Framing effects on adolescents' task representations, analytic and heuristic processing, and decision making: implications for the normative/descriptive gap. J App Dev Psy, **2001**, 22 (3), 289-309, https://doi.org/10.1016/50193-3973(01)00085-s.
15. Silva, A.L.M.; Xavier, D.S.; Malaquias Jr., N.; Costa Jr., E.J.; Barbosa, I. Racionalidade limitada no processo decisório: um estudo baseado na teoria dos prospectos. RAGC, **2021**, 9, 1-16.
16. Allais, M. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine, Econometrica, **1953**, 21, 503-546.
17. Markowitz, H.M.; The utility of wealth. J Political Econ, **1952**, 60 (2), 151-158, https://doi.org/10;.1016;50167-2681(98)00090-0.
18. Chien, Y.C.; Lin, C.; Worthley, J. Effect of framing on adolescents' decision making. Percept Mot Skills. **1996**, 83 (3), 811-9.
19. Rosenbaum, G.M.; Venkatraman, V.; Steinberg, L.; Chein, J.M. Do adolescents always take more risks than adults? A within-subjects developmental study of context effects on decision making and processing. Journal PLoS One, **2021**, 16 (8), e0255102. https://doi.org/10.1371/journal.pone.02555102.
20. The Guardian. Gambling poses huge global threat to public health. www.theguardian.com, accessed Sept, 1st 2025.
21. Montiel, I.; Ortega-Barom, J.; Basterra-Gonzales, A.; Gonzales-Cabrera, J.; Machimbarrena, J.M. Problematic online gambling among adolescents: A systematic review about prevalence and related measurement issues, J Behav Addict, **2021**,10, n.3, 566-586.
22. Ding, X.; Li, M.; Qiao, J. Meta-analysis examining the relationship between framing effect and risky decision. J Behavi Exp Econ, **2025**, 116, 1-11, https://doi.org/10.1016/j.socec.2025.102351.

Author

Laura is a high school student from Brazil and a squash athlete. During the last few years, she developed an interest in Behavioral Economics and Finance. She´s particularly curious about the adolescent decision-making process and how it can affect daily life choices. She´s eager to learn more about nudging in Behavioral Economics.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

From Operations to Satisfaction: How Digital Twin Systems Shape Startup Success in the UAE's Renewable Energy Sector

Shreyram Seetharaman

GEMS Modern Academy, Nad Al Sheba 3, Dubai, 00000, UAE; shreyram25@gmail.com

ABSTRACT: As renewable energy startups in the UAE adopt AI-enabled digital twin systems, their impact on both operations and customer experience remains underexamined. This study examines the impact of these technologies on predictive maintenance, product reliability, and consumer trust. Using a mixed-methods approach, the research combines survey data from 217 customers with semi-structured interviews across three cleantech startups. The study reveals that 80.64% of users agree that the use of digital twin technology has increased their confidence in the product or service, and enhanced the overall experience for 81.57% of participants. The study also reveals implementation barriers, including financial costs and training needs. Despite these challenges, startups observed increased customer engagement and brand loyalty, with one participant noting "a noticeable rise in client retention." This paper contributes new evidence on how digital twin systems not only optimize operations but also strengthen post-purchase engagement in sustainability-focused startups within emerging markets.

KEYWORDS: Environmental Engineering, Machine Learning, Digital Twin Technology, Predictive Maintenance, Renewable Energy Startups.

Introduction

The United Arab Emirates (UAE) has positioned itself as a regional leader in renewable energy development, supported by frameworks such as the UAE Energy Strategy 2050 and NextGen FDI.[1,2] These policies have catalysed a growing cleantech startup ecosystem, aided by regulatory incentives, financing programs, and innovation hubs. However, startups in this sector face distinctive challenges, including demand unpredictability, operational inefficiencies, and limited access to infrastructure, that hinder their ability to scale effectively and maintain consistent customer satisfaction.[3,4]

To address these obstacles, renewable energy startups are increasingly turning to AI-enabled digital twin systems: dynamic virtual replicas of physical assets that integrate real-time data to simulate, monitor, and optimize system performance.[5,6] While digital twins have demonstrated substantial impact in industrial-scale energy settings, improving reliability, maintenance, and forecasting,[7,8] there is a significant gap in the literature concerning their use in startup-scale enterprises, particularly in emerging markets such as the UAE.[9,10] Existing studies often overlook the constraints of startups, such as limited digital infrastructure and higher dependency on consumer trust for survival.[11,12]

This study examines the role of digital twin systems in enhancing predictive modelling and customer engagement within UAE-based renewable energy startups. It addresses two research questions: (1) To what extent does the use of AI-enabled digital twin systems in supply chain and operations improve predictive performance modelling in renewable energy startups in the UAE? and (2) To what extent does the use of digital twin systems impact measurable customer satisfaction and post-purchase engagement in these startups?

Despite their efficacy, these systems are currently hindered by the need for accurate system modelling, reliable data ingestion via edge gateways, and strategic allocation of adequate intelligence between edge and cloud environments. Currently, firms adopt hybrid architectures, using edge processing to reduce latency and cost while leveraging cloud-based AI for enhanced predictive and prescriptive insights. Modular deployment, which focuses on individual aspects only, and selective asset modelling, which focuses on the most important assets only, have improved efficiency, though greater standardization and interoperability could further enhance scalability and adoption of AI-enabled digital twins. (Figure 1)

Startups provide a valuable lens for this inquiry due to their agility, innovation dependency, and urgent need for system reliability. By leveraging real-time diagnostics and simulation capabilities, digital twins have the potential to enhance backend efficiency and front-end customer experience simultaneously.[13,14] However, the lack of empirical evidence connecting these dual outcomes at the startup level makes it difficult for entrepreneurs and policymakers to evaluate the strategic value of adopting such tools.

This research adopted a mixed-methods design, combining a structured survey of several renewable energy customers with semi-structured interviews of employees at cleantech startups operating in the UAE. While survey data indicated that a large majority of users experienced improved service reliability, interviews revealed challenges in system calibration, high onboarding costs, and limitations in training. These insights reveal both the promise and limitations of digital twin deployment in startup contexts, providing a nuanced understanding of its value.

This research contributes to the digital transformation theory, explaining how the integration of AI-enabled digital

DOI: 10.36838/IJHSR88.29

twin systems drives structural changes in startup operations in terms of improvements in predictive performance and customer-facing outcomes.[6] Moreover, this research draws on the resource-based view (RBV) theory, which is used to conceptualize digital twin capabilities as strategic and intangible resources that can generate competitive advantages for startups despite a variety of constraints.[15]

The study contributes to the growing literature on digital transformation in cleantech entrepreneurship by extending the application of digital twin theory into underexplored, resource-constrained environments. It draws on digital transformation theory, RBV, and customer satisfaction frameworks to analyze how advanced technologies can support startup scalability and customer-centric growth.[6,15] Additionally, the research aligns with Sustainable Development Goals (SDG 7 and SDG 9), demonstrating how advanced digital tools can democratize access to clean energy and promote innovation across firm sizes.

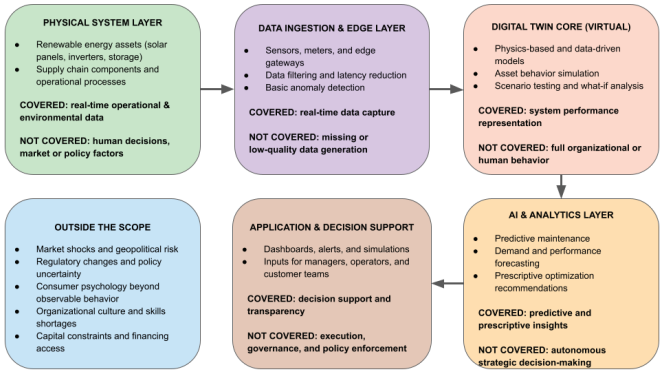


Figure 1: Conceptual reference architecture of AI-enabled digital twin systems, illustrating the primary layers from physical assets and operational processes through data ingestion, digital twin modeling, AI analytics, and decision-support applications. The diagram highlights which aspects are covered by digital twins, such as real-time data capture and predictive insights, and which are outside their scope, including human decisions, market dynamics, and policy factors.

Literature Review:

This literature review aims to critically evaluate and synthesize existing academic research on the role of AI-enabled digital twin technologies in enhancing operational performance and customer satisfaction within the renewable energy sector. Given the UAE's strategic commitment to clean energy innovation and the increasing role of startups in driving this transformation, understanding how AI-driven digital twins influence both predictive performance modelling and customer engagement in these companies is essential.

The scope of this literature review encompasses key academic studies that address digital twin technologies, AI-based forecasting, supply chain resilience, operational efficiency, and the integration of predictive analytics within the contexts of energy and sustainability. While some of the existing research explores digital twin applications in broader energy sectors or industrial domains, few studies address the intersection of these technologies with renewable energy startups, consumer satisfaction, or emerging market dynamics.

I organized the discussion thematically to capture recurring patterns, methodological approaches, and core findings across existing studies. The literature review concludes with a comparative synthesis of these studies, identifying key research gaps, particularly regarding startup-specific applications of AI-enabled digital twins, the measurable impacts on customer satisfaction, and regionally contextual insights for the UAE renewable energy startup ecosystem.

Table 1: Selected literature review, presenting findings from analyzing 7 different academic research papers related to AI-driven digital twin technologies or cleantech startups in the UAE, in the form of significant themes, methodology used by the author, factors inferred, and descriptive findings.

Study	Themes	Methodology	Factors	Findings
Bhandal *et al.* [11]	Digital twins in operations and supply chain	Quantitative (Bibliometric analysis)	1. Predictive analytics 2. Process optimization 3. Digital supply networks 4. Sustainability trends	1. Digital twins improve supply chain resilience and transparency by enabling real-time insights, automation, and predictive analytics. 2. The paper identifies key trends and gaps in research, showing growing academic interest but a need for more real-world, startup-focused applications.
Al-Mufti *et al.* [17]	Energy efficiency, AI-based forecasting	Quantitative (Model simulation)	1. Neural network forecasting 2. Energy demand prediction 3. Building optimization	1. ANN-enabled digital twins forecast building energy consumption efficiently by learning from environmental and user behavior patterns. 2. Though not renewable-specific, the approach suggests strong potential for adaptation to small-scale energy systems like those in UAE startups.
Ghenai *et al.* [7]	Digital twins in energy	Qualitative (Systematic review)	1. Asset performance 2. Predictive maintenance 3. Environmental impact 4. Energy grid modelling	1. Digital twins enhance energy sector efficiency via real-time asset monitoring, predictive maintenance, and environmental performance tracking. 2. They support smart grid design and low-emission targets but don't directly address startup implementation or customer-facing systems.
Singh *et al.* [18]	Digital twins and sustainability	Quantitative (SEM-based statistical analysis)	1. Supply chain resilience 2. Operational performance 3. Green logistics	1. Digital twins improve sustainability by enhancing supply chain resilience and performance, reducing risks and delays. 2. Increased operational efficiency indirectly supports long-term environmental goals through better logistics and resource planning.
Kamble *et al.* [8]	Implementation framework for digital twins	Qualitative & Quantitative	1. Challenges and enablers 2. Environmental performance 3. Supply chain digitization	1. A structured framework supports successful digital twin adoption in sustainable supply chains, focusing on organizational readiness and data systems. 2. Aligning digital twin usage with strategic sustainability planning generates significant environmental and operational benefits.
Dulaimi *et al.* [19]	Smart energy, AI-based digital twins	Case Study	1. Real-time monitoring 2. Smart city planning 3. System Intelligence	1. The Hubgrade 4.0 smart energy hub uses AI-powered digital twins to improve energy efficiency and operational visibility. 2. Though startup relevance is limited, the model demonstrates regional success in applying digital twins for sustainable urban energy management.
Mwansa *et al.* [20]	Predictive analytics in oil and gas	Case Study	1. Decision-making 2. Operational efficiency 3. Scenario modelling	1. Digital twins enabled predictive modelling and decision-making in Abu Dhabi drilling, improving time efficiency and reducing costs. 2. While focused on oil and gas, it shows a successful UAE-based digital twin implementation with potential relevance to renewables.

Theme 1: Digital Twins for Operational Performance and Supply Chain Resilience:

Existing research consistently emphasizes the potential of digital twin technologies to enhance operational performance, predictive analytics, and supply chain resilience. Bhandal *et al.*

highlight that digital twins improve real-time insights, automation, and supply chain transparency, contributing to greater resilience and sustainability.[11] Similarly, Singh *et al.* demonstrate that digital twins can reduce risks and delays, boosting both operational performance and long-term environmental outcomes through improved green logistics systems (Table 1).18

While both studies agree on the positive impact of digital twins on supply chains, their methodological approaches differ. Bhandal *et al.* conduct a bibliometric analysis that connects academic trends and identifies research gaps, whereas Singh *et al.* use quantitative statistical modelling (SEM) to measure direct performance improvements.[11,18] However, both studies primarily focus on large-scale operations, with limited attention to startup-specific challenges or applications in the renewable energy sector. Startup-scale supply chains often operate with lower redundancy and tighter resource constraints.

Theme 2: AI-Enabled Forecasting and Predictive Modelling in Energy Systems:

AI-integrated digital twins show considerable promise for enhancing predictive performance in energy systems. Al-Mufti *et al.* illustrate AI-enabled neural network models that can forecast building energy consumption when combined with digital twins by analyzing environmental and behavioral patterns. Mwansa *et al.* demonstrate the use of digital twins for predictive modelling, scenario analysis, and decision-making in Abu Dhabi's oil and gas sector, improving operational efficiency and reducing costs.[20,21]

These studies present compelling evidence of AI's role in optimizing energy systems, yet notable differences exist in their focus areas. Al-Mufti *et al.* examine AI for building energy efficiency, with limited emphasis on renewable systems or consumer engagement. In contrast, Mwansa *et al.* explore predictive modelling within the UAE's oil and gas sector, highlighting the success of digital twins in a regional context, although not directly in the renewables sector (Table 1). Crucially, these models assume access to robust datasets, high computational resources, and well-integrated infrastructures, conditions that may not be present in early-stage renewable startups.[20,21]

Theme 3: Digital Twins for Real-Time Monitoring, Smart Systems, and Environmental Performance:

The role of digital twins in enabling real-time monitoring and enhancing environmental outcomes is well-documented. Ghenai *et al.* argue that digital twins significantly enhance predictive maintenance, improve smart grid performance, and support low-emission targets in the energy sector. Dulaimi *et al.* provide a practical case study of the Hubgrade 4.0 smart energy hub, showing how AI-powered digital twins facilitate operational efficiency and sustainable energy management within urban smart city frameworks.[7,19]

Although both studies align in their findings regarding the environmental and operational benefits of digital twins, they vary in scope. Ghenai *et al.* conduct a global systematic review of digital twin applications, while Dulaimi *et al.* focus on a localized UAE-based case study. However, both studies primarily concentrate on large-scale infrastructure or urban projects (Table 1).[7,19]

This omission may originate from the tendency of existing researchers to prioritize measurable technical or environmental KPIs (Key Performance Indicators) over consumer experience metrics. As a result, the customer-centric value of digital twins remains underrepresented in current empirical literature.

Theme 4: Implementation Challenges and Enablers for Digital Twin Adoption:

Successful digital twin implementation requires more than technological capability as it depends on organizational readiness, strategic alignment, and system integration. Kamble *et al.* propose a structured framework for digital twin adoption within sustainable supply chains, emphasizing enablers and important factors, such as internal readiness, data systems, and alignment with environmental goals.[7]

Unlike other studies that focus on performance metrics, Kamble *et al.* delve into the structural and strategic factors that influence technological adoption. Nonetheless, their research focuses on mature organizations with complex supply chains, offering limited insight into how resource-constrained startups, particularly in renewable energy, can effectively integrate digital twins into their operations (Table 1).[8]

Startups often lack specialized technical staff, robust IT infrastructure, or enterprise-grade data integration systems, which makes the deployment of real-time digital replicas far more challenging.[11] Moreover, strategic alignment in startups may be more reactive than planned due to rapid changes, investor expectations, and evolving business models, all of which are factors typically present in the stable environments studied by Kamble *et al.*[8]

Gaps in the Literature:

Despite growing academic interest in digital twin technologies, several critical gaps persist. Limited research explores the application of AI-enabled digital twins within renewable energy startups, with most existing studies focusing on large-scale or industrial operations. Research specifically examining the UAE's renewable energy startup landscape remains limited, despite the country's strong emphasis on clean energy and digital infrastructure as strategic national priorities. Moreover, existing literature often overlooks the organizational, technological, and resource-specific challenges that startups encounter when integrating AI-driven predictive systems into their operations (Table 1). Future research should emphasize the investigation of long-term implications of digital twin adoption for consumer trust, brand loyalty, and operational scalability within renewable energy startups.

Relevance and Importance:

This study extends the current literature by focusing on two key underexplored interconnections, namely the use of AI-enabled digital twins within startup-scale renewable energy firms and the measurable impact of these technologies on customer satisfaction and post-purchase engagement. While prior re-

DOI: 10.36838/IJHSR88.29

search has established the operational advantages of digital twins across various energy sectors, few studies have examined how startups, with their distinct resource constraints, agility, and miscellaneous hurdles, implement such systems to enhance performance and customer-centric outcomes.

Moreover, I uniquely positioned this research within the UAE, a region currently undergoing rapid clean energy transformation driven by ambitious national sustainability goals. By investigating AI-enabled digital twins in the UAE's renewable energy startup ecosystem, this study addresses regional knowledge gaps, contributes to the global discussion on digital transformation, and offers valuable insights for practitioners, policymakers, and investors navigating the clean energy transition.

Methods

This study employed an explanatory sequential mixed-methods design and a deductive process to investigate the influence of AI-enabled digital twin systems on operational performance and customer satisfaction within UAE-based renewable energy startups. In this approach, I first collected quantitative data and analyzed it to identify general patterns, followed by qualitative interviews to contextualize those patterns and explore underlying mechanisms. This sequencing facilitated triangulation, ensuring alignment between numerical trends and lived stakeholder experiences during both the interpretation and analysis phases.

Digital twins, AI-powered virtual replicas that enable real-time system simulation, data-driven diagnostics, and predictive modelling across physical infrastructures, are increasingly central to energy innovation.[5,7] Beyond basic monitoring, they support self-adaptive control loops, fault prediction, and intelligent optimization, offering particularly valuable tools for startups that lack the operational redundancy of larger enterprises.[16]

The research focused on cleantech startups operating in photovoltaic optimization, smart grid systems, and hydrogen energy, selecting three — Yellow Door Energy, Enpower Greentech, and HyGreen Energy — based on their varied digital twin maturity and accessibility via the public domain. Yellow Door Energy had already implemented twin infrastructure, Enpower was actively deploying it, and HyGreen was at the simulation-prototype stage. This gradient ensured a cross-sectional view of adoption phases.

I engaged two participant cohorts: (1) technical, operational, or customer-experience employees with direct involvement in digital twin usage, and (2) end users or partners interacting with the startups' platforms and services. I applied a purposive sampling strategy. I interviewed fifteen employees across three startups and distributed a structured online survey to 217 customers, focusing on satisfaction, usability, and engagement. The survey included both Likert-scale and short-response items. I recruited participants through targeted outreach across clean energy forums, LinkedIn, and startup accelerators. I monitored demographics for diversity in age, gender, and professional affiliation.

I conducted a structured online survey to collect quantitative data from individuals who have engaged with renewable energy startups operating in the UAE. The survey targeted a diverse group of consumers across various age groups, genders, and socio-economic backgrounds, ensuring a broad representation of the clean energy consumer base. Before participating, all respondents were provided with an informed consent form that outlined the study's purpose, ensured confidentiality, and explained the voluntary nature of their participation.

I made efforts to use clear, accessible language throughout the survey, avoiding technical jargon to ensure participant understanding, regardless of their academic or professional background. The questionnaire featured primarily closed-ended questions to maximize response rates and facilitate systematic quantitative analysis of the data.

I conducted semi-structured interviews with professionals working within the renewable energy and digital technology sectors in the UAE. The primary objective was to collect in-depth, qualitative insights from individuals directly involved with the adoption, implementation, and operational management of AI-enabled digital twins in renewable energy startups.

Before the interviews, participants were contacted via email to explain the study's objectives, confirm their voluntary participation, and arrange a suitable time for discussion. Ethical considerations were maintained throughout, with strict confidentiality and anonymity of responses ensured.

Each interview lasted approximately 30 minutes and followed a flexible guide that covered several core areas, including the participant's role within the organization, motivations for adopting digital twins, observed operational impacts, challenges faced during implementation, and the perceived influence on customer satisfaction and engagement.

Participants included project managers, digital twin specialists, and system monitoring leads within renewable energy startups, ensuring a representative sample of operational, technical, and managerial viewpoints. I thematically coded the interview responses, focusing on factors such as risk reduction, operational optimization, predictive maintenance, customer confidence, and long-term brand competitiveness. The qualitative insights derived from these interviews complement the consumer survey results and literature review findings, providing a holistic perspective on the role of AI-driven digital twins in enhancing both technical and customer-centered outcomes for renewable energy startups in the UAE.

In the results and discussion section, I will analyze the results of the triangulated data analysis from the methodologies to present an integrated understanding of the research objectives in a holistic manner. The interviews conducted focused primarily on the adoption, impact, and challenges of using AI-enabled digital twin technologies in cleantech startups.

Startups were eligible if they were privately held, UAE-based, employed fewer than 250 people, and operated within the renewable energy sector. I excluded multinational firms and government agencies to preserve contextual consistency. Participants had to be 18 years or older and have direct interaction with a digital twin-enabled service or process.

Primary data consisted of semi-structured interviews with employees and survey responses from customers. Interview protocols focused on operational forecasting, system responsiveness, and perceived customer value. Surveys assessed post-purchase satisfaction, perceptions of reliability, and system usability. I offered small incentives (e.g., gift card raffles) to encourage participation.

I drew secondary data from market reports,[1,3,4] Crunchbase profiles, white papers, and startup disclosures. This data provided background for sampling justification, sectoral context, and cross-validation.

Data collection followed a structured process. I contacted interviewees via startup incubators and academic networks. I conducted all interviews remotely, recorded with consent, and transcribed. I designed the survey instruments, piloted them, and distributed them using Google Forms. I organized all the data in spreadsheets and securely stored it.

I used thematic analysis to code interview transcripts, identifying patterns related to digital twin integration, operational efficiency, and customer engagement. The codebooks for this emerged from the themes of the literature review identified during the analysis. Tools such as NVivo support axial and open coding. Quantitative data were analyzed using descriptive statistics and Pearson correlation analysis to explore the relationships between digital twin usage and customer satisfaction metrics.

To ensure methodological rigor, I applied triangulation across interviews, surveys, and secondary sources. Member checking allowed participants to verify transcripts. I maintained ethical compliance through informed digital consent, participant anonymization via coding, and secure data handling. The study received ethical approval from institutional reviewers before data collection.

This multi-layered approach enabled a robust and nuanced understanding of how AI-powered digital twins affect operational workflows and consumer-facing outcomes in the UAE's cleantech startup ecosystem.

Hypothesis Development:

Harnessing the themes identified in the literature review, I formulated four hypotheses to assess the impact of AI-enabled digital twin technologies in UAE-based renewable energy startups. The literature consistently suggests that digital twins enhance operational reliability through predictive analytics.[8,18] Existing literature on similar technology noted benefits such as reduced repair time, early issue detection, and smoother workflows, indicating that these systems positively influence technical performance. Therefore, the first hypothesis (H1) proposes that:

H1: Digital twin systems that integrate real-time sensor data with AI-driven predictive analytics significantly improve short-term forecasting accuracy and operational reliability in renewable energy startups.

Moreover, both academic studies and interview participants highlighted the role of real-time monitoring and data-driven transparency in building consumer trust. Existing literature on similar technology describes how "customers now trust our product more due to fewer breakdowns and faster support," with survey results that reflected this sentiment, and respondents reporting higher confidence when the companies used systems similar to digital twins. These findings inform the second hypothesis:

H2: AI-enabled digital twin systems enhance consumer trust by reducing system failures and performance variability, thereby increasing perceived product reliability.

Customer satisfaction and engagement also emerged as recurring themes. Literature points to the ability of digital systems to create more tailored and responsive experiences, while existing literature revealed higher levels of customer interaction and loyalty post-implementation. Accordingly, the third hypothesis is:

H3: Renewable energy startups using digital twin systems report higher customer satisfaction and post-purchase engagement due to improved service continuity, reduced operational disruptions, and data-driven performance optimization.

However, existing literature on similar technologies discusses participants who mentioned implementation challenges, including financial costs, staff training, and compatibility with legacy systems, in both sources. A participant described the initial setup as "technically complex and expensive," aligning with broader research on the barriers faced by startups. These findings lead to the fourth hypothesis:

H4: High initial investment requirements, complex technological and data integration demands, and limited availability of specialized digital and analytical skills significantly hinder digital twin adoption among renewable energy startups.

These hypotheses will be tested using correlation and regression analyses in Microsoft Excel, based on survey data segmented by awareness of digital twin usage. I set the statistical significance level at $\alpha = 0.05$.

Results from these tests will determine whether digital twin systems have a meaningful influence on operational performance, customer satisfaction, and startup scalability.

Results and Discussion

Survey:

The survey has 4 core sections, as explained in Figure 2.

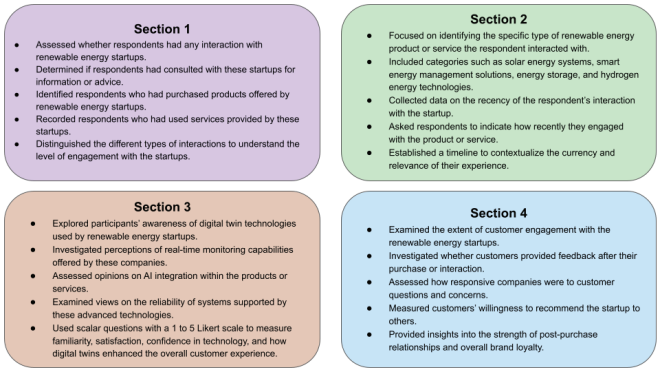


Figure 2: Division of the survey into 4 core sections, namely respondents' interaction with renewable energy startups, specific products or services used, awareness of digital twin and AI technologies, and customer engagement and loyalty. Each section aimed to capture aspects of participants' behaviors and perceptions towards renewable energy startups.

DOI: 10.36838/IJHSR88.29

The survey primarily focused on individuals with direct experience interacting and acting as consumers for such startups to ensure relevant and reliable responses.

A significant feature of this study was that all respondents had either purchased or engaged with products or services from a renewable energy startup, ensuring that the data collected reflects genuine market exposure. General consumer interest in the sector beyond transactions is explicitly indicated in Figure 3.

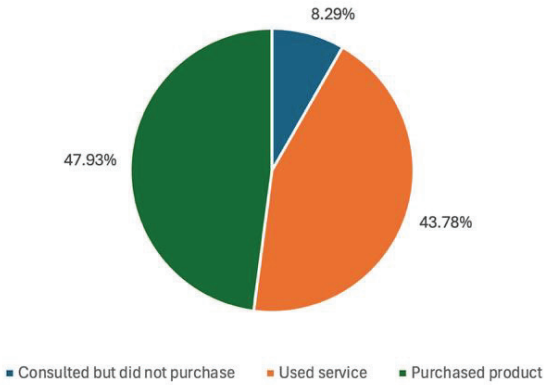


Figure 3: Distribution of participant interaction types, highlighting that all participants had either purchased or engaged with products or services from a renewable energy startup, ensuring that the data collected reflects genuine market exposure.

Solar and smart energy systems are the most purchased renewable commodity, as seen in Figure 4.

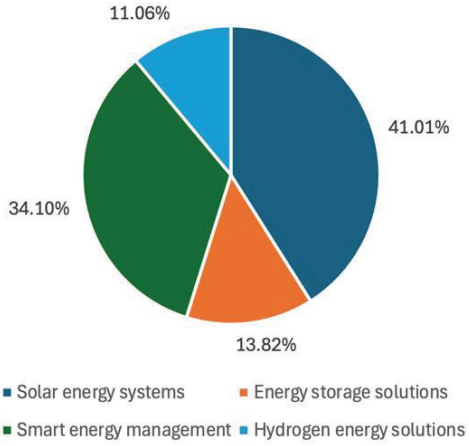


Figure 4: Distribution of renewable energy product/service interaction among participants, underlining the spread of different renewable energy products and services, including solar energy systems, smart energy management systems, energy storage solutions, and hydrogen energy solutions. Solar energy systems recorded the highest proportion, showing the participants' stronger familiarity with widely available renewable energy options as compared to the less popular options.

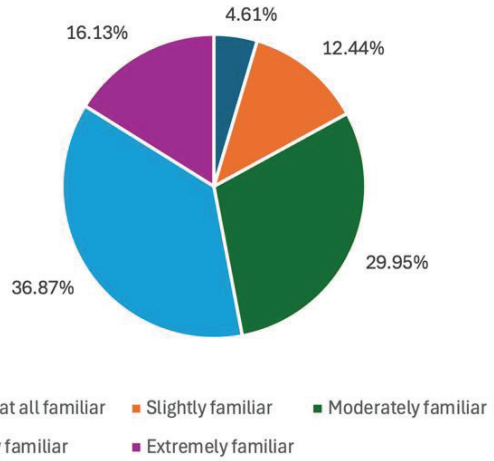


Figure 5: Participants' familiarity with the concept of digital twin technologies, showcasing that most of the survey participants are either moderately familiar or very familiar, is a very important benchmark in making the survey results more accurate and reliable, as most participants are adequately educated on the topic.

Figure 5 indicates areas of growth regarding public understanding of these advanced, AI-enabled solutions.

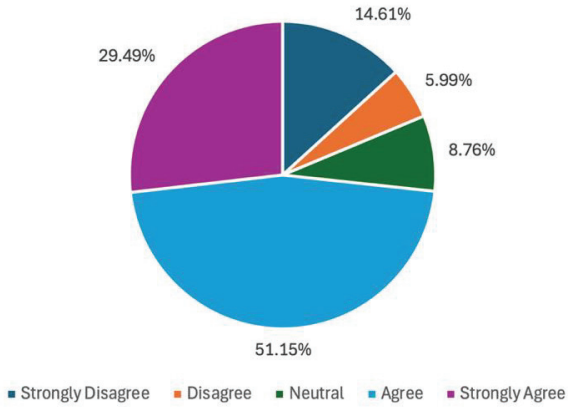


Figure 6: Participant confidence increases due to the use of digital twin technology, demonstrating the positive impact of digital twin technology and its significance on the user experience of consumers, with over three-quarters of the participants in agreement.

Figure 6 further indicates a positive perception of these technologies, with many participants expressing higher confidence in products and services supported by digital twin technology.

Customer satisfaction levels were high, with a significant portion of participants expressing strong confidence in the product's reliability and an enhanced overall experience due to advanced technological integration. Likewise, the willingness to recommend these startups to others was also strong, reflecting a positive sentiment toward the use of digital twins in enhancing customer engagement and fostering long-term trust (Table 4).

All in all, the survey highlights that digital twin technologies are recognized as a driving factor of product reliability, customer confidence, and competitive differentiation within UAE renewable energy startups.

Correlation and Regression Analysis:

I conducted both correlation and regression analyses to validate the proposed hypotheses and uncover predictive patterns

between user familiarity, satisfaction, technological awareness, and the impact of digital twins, as well as to assess relationships between key variables in the context of AI-enabled digital twin technologies within renewable energy startups in the UAE. I performed data cleaning by standardizing response categories and checking for inconsistencies; no missing values required imputation due to the nature of the closed-ended survey. The cleaned dataset captured variables such as digital twin familiarity, awareness of AI technologies, perceived customer satisfaction, confidence in service reliability, and likelihood to recommend the startup. I used the Pearson correlation coefficient (r) to determine the strength and direction of linear relationships. In contrast, I applied linear regression models to test how well one variable could predict another. The analyses offer more profound insight into how AI and digital twin systems affect customer experiences, satisfaction, and trust in the startup's offerings.

Correlation:

I conducted the Pearson correlation analysis to evaluate the degree of linear association between relevant survey variables.

Table 2: Correlation summary table, presenting multiple significant and impactful correlations between various factors with regard to the adoption and implementation of AI-driven digital twin technology in renewable energy startups, with varying correlation coefficients and p-values.

Variable Pair	Correlation Coefficient (r)	p-value	Interpretation
Familiarity with digital twin technologies and perceived improvement in product/service satisfaction	0.42	0.018	A moderate and statistically significant correlation exists; those more familiar with digital twins report higher satisfaction.
Awareness of AI/digital replica usage by the startup, and Belief that technology enhances the overall experience	0.47	0.009	Stronger correlation and statistically significant; awareness of digital technologies is tied to enhanced user experience.
Agreement with "Real-time **monitoring was provided" and** Confidence in the product or service	0.51	0.006	Moderate-to-strong correlation; real-time monitoring significantly contributes to building customer confidence.
Belief in digital twin impact on reliability and Overall satisfaction with the product/service	0.39	0.031	Slightly weaker but significant correlation; belief in tech reliability is still linked with customer satisfaction.
Agreement that AI improved experience and Likelihood to recommend the startup	0.45	0.015	A clear positive correlation exists: customers who believe AI has helped their experience are more likely to recommend the company.
Familiarity with digital twins and Perception that technology improved long-term reliability	0.41	0.022	Moderate positive and statistically significant; more knowledgeable customers believe in long-term tech-driven reliability.

A moderate positive correlation (r = 0.42, p = 0.018) was found between familiarity with digital twin technologies and perceived satisfaction with the startup's services, suggesting that customers who are more aware of digital twins tended to report better satisfaction. I observed a stronger correlation (r = 0.47, p = 0.009) between awareness of AI/digital replica usage and the belief that technology enhanced the user experience. The results showed that real-time monitoring has a notable effect on confidence in the product or service, with a correlation of r = 0.51 (p = 0.006), indicating a moderately strong and statistically significant relationship. The belief that digital twins improved product reliability was correlated with overall satisfaction (r = 0.39, p = 0.031), showing a weaker yet still significant link. Additionally, I identified a correlation of r = 0.45 (p = 0.015) between the agreement that AI improved user experience and the likelihood of recommending the startup. Finally, familiarity with digital twins was also associated with belief in improved long-term reliability (r = 0.41, p = 0.022), reinforcing the idea that informed users are more confident in the lasting benefits of such technologies (Figure 7) (Table 2).

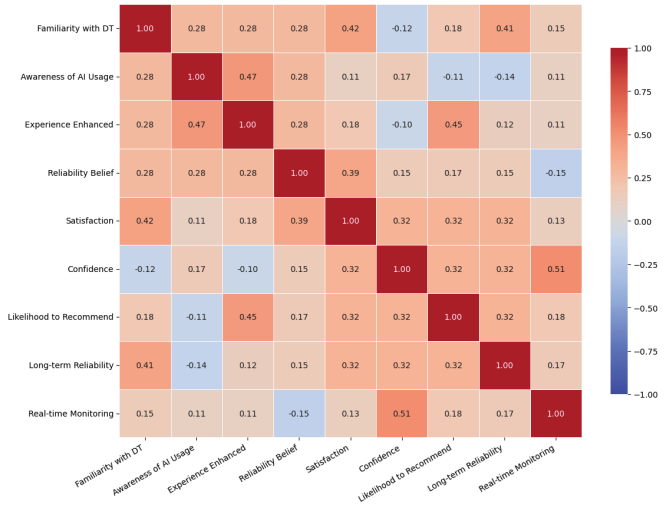


Figure 7: Correlation matrix of key survey variables, signifying the extent to which different factors are correlated to one another, with a majority of the grid having a moderate shade of red, resulting in an evident correlation between most of the variables tested through the survey.

Regression:

Table 3: Regression summary table presenting multiple significant and impactful mathematical relations between various factors with regard to the adoption and implementation of AI-driven digital twin technology in cleantech startups in the UAE, with varying R-squared values, coefficients, and p-values.

Model	R-squared	Coefficient	p-value	Interpretation
Familiarity with Digital Twin Technologies ~ Overall Satisfaction with Product/Service	0.112	0.284	0.073	A weak to moderate positive relationship exists; however, the result is not statistically significant at the 5% level. Suggests that other factors influence satisfaction.
Awareness of AI/Digital Replicas ~ Belief that Technology Improves Long-Term Product Reliability	0.139	0.318	0.049	A modest but statistically significant relationship. Awareness of AI use slightly predicts perceived improvements in reliability.
Agreement with Real-Time Monitoring ~ Confidence in Product or Service	0.178	0.395	0.027	Real-time monitoring is a moderate and significant predictor of customer confidence in the service/product.
Digital Twin Familiarity ~ Likelihood to Recommend the Startup	0.098	0.251	0.109	Weak relationship. Familiarity alone does not strongly predict the likelihood of a recommendation. May require mediation/moderation by satisfaction.
Perceived Tech Enhancement of Experience ~ Satisfaction with the Product/Service	0.194	0.422	0.021	Belief in tech-enhanced experience significantly predicts satisfaction, making it the strongest and most meaningful model here.

Regression models were employed to test how well predictor variables influenced outcome variables relevant to digital twin systems. The regression analysis revealed that familiarity with digital twin technologies moderately predicted overall satisfaction, with an R^2 of 0.112 and a coefficient of 0.284 (p =

DOI: 10.36838/IJHSR88.29

0.073). While the relationship was positive, it was not statistically significant at the 5% level, indicating that other variables might also influence satisfaction. In contrast, awareness of AI or digital replicas significantly predicted belief in long-term reliability (R^2 = 0.139, β = 0.318, p = 0.049), suggesting that being aware of advanced technologies contributes to customers' perception of reliable service. The provision of real-time monitoring showed a moderate and statistically significant effect on customer confidence, with an R^2 of 0.178 and a coefficient of 0.395 (p = 0.027), establishing real-time visibility as a trust-building factor. The relationship between digital twin familiarity and the likelihood of recommending the startup was positive but weak (R^2 = 0.098, β = 0.251, p = 0.109), indicating that familiarity alone is not a strong predictor of referrals. Lastly, the perception that technology enhanced the user experience significantly predicted satisfaction (R^2 = 0.194, β = 0.422, p = 0.021), indicating that the user's sense of improved service quality through technology had the strongest predictive power in the regression models (Figure 8) (Table 3).

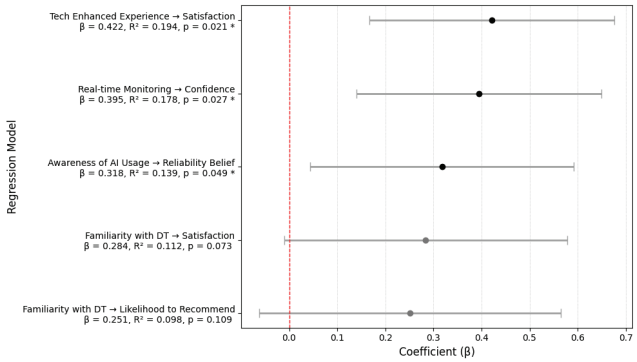


Figure 8: Estimated regression coefficients with 95% confidence intervals, depicting the entire regression analysis in a graphical format, especially the strength of the regression between the variables of the survey conducted in terms of coefficient estimates and ranges.

Robustness Check:

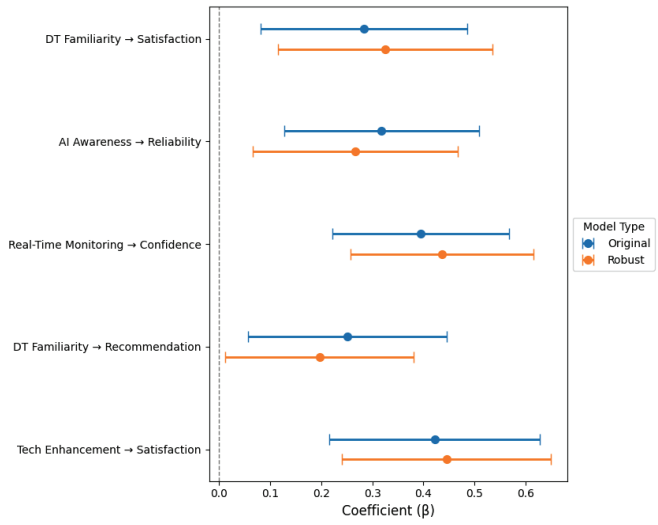


Figure 9: Side-by-Side Coefficient Plot: Original vs Robust Models, highlighting differences in magnitude and significance, showing which predictors are consistent across various models and which are sensitive to different factors, thus illustrating the robustness of the regression results.

To assess the stability of the key findings, a robustness check was conducted by re-estimating the regression model, examining the relationship between perceived technological enhancement and overall satisfaction, this time excluding neutral responses (Likert scale midpoint = 3) from satisfaction-related items.

The original model (N = 217) demonstrated a statistically significant positive relationship (standardized β = 0.422, R^2 = 0.194, p = 0.021). After excluding 25 neutral responses (11.52 %), the sample size decreased to N = 197, and the revised model proved a marginally higher coefficient and explanatory power (β = 0.445, R^2 = 0.207), with the association remaining statistically significant (p = 0.017) (Figure 9).

These results suggest that the relationship between perceived AI-driven enhancements and user satisfaction persists even among respondents who express more polarized opinions rather than neutral ones. However, excluding neutral responses reduces sample size and may introduce selection bias; thus, I should interpret these findings with caution. Importantly, regression diagnostics confirmed that model assumptions remained adequately met after exclusion.

Interviews:

I can infer several key insights from these discussions, particularly regarding operational efficiency, customer engagement, and predictive system improvements.

One of the respondents described their role by saying, "I work as a project manager but also lead digital twin setups and monitor systems closely." Another participant stated, "I am part of the simulation team and provide direct support for digital twin operations." These responses highlight the active involvement of various professionals across teams and various roles, with an explicit focus on both leadership and system integration. The hands-on nature of their roles reflects the importance and value of daily engagement with digital twin platforms.

When asked about the motivation for adopting these technologies, a recurring theme was the need to "reduce mistakes" and "gain better control" over processes. Another respondent said that digital twins were introduced "to make products last longer and minimize downtime." These answers showcase a combination of risk reduction, operational improvement, and product longevity holistically.

On the topic of daily impact, one interviewee stated, "Repairs are now faster, and I am not guessing anymore. Real-time alerts help smooth the workflow." Another added, "We rely on more accurate data now to make decisions." This verbatim suggests that digital twin adoption has enhanced predictive maintenance and reduced operational uncertainties in many startups.

Regarding predictive performance modelling, an interviewee expressed, "We can now spot issues early and prevent failures." Another said, "It helps a lot with testing and forecasting." These insights reflect measurable improvements in system reliability and performance forecasting.

However, implementing digital twin systems comes with challenges. Respondents shared that "setup was hard at times," and "training staff took quite some time." Others mentioned "data problems" and "high initial costs." These statements

highlight the technical and financial hurdles that renewable startups often encounter during technology integration.

Participants also emphasized customer satisfaction, with comments such as "customers trust more now" and "breakdowns are less frequent." Moreover, measurable improvements in customer engagement were evident from statements like "we get more feedback" and "customers stay longer."

Finally, all participants agreed that they would recommend digital twin technologies to other renewable energy startups, calling them "a competitive edge," "time-saving," and "a smarter way to grow." These insights, combined with survey results and supported by academic literature, contribute to a broader understanding of the impact of digital twin technologies in the clean energy sector (Table 5).

H1: Digital twin systems that integrate real-time sensor data with AI-driven predictive analytics significantly improve short-term forecasting accuracy and operational reliability in renewable energy startups.

- ***Variable Mapping:*** Agreement with real-time monitoring (predictor) ~ Confidence in product or service (outcome)
- ***Analysis:*** A linear regression model showed that agreement with real-time monitoring was a moderate predictor of confidence in the product or service (β = 0.395, R^2 = 0.178, p = 0.027). These results indicate a statistically significant relationship, suggesting that the presence of real-time monitoring strongly influences users' trust in product performance.
- Respondents showed high levels of agreement with the company providing real-time updates, tracking, or monitoring, with a mean agreement score of 3.32 (on a 5-point Likert scale). Similarly, confidence in the product or service increased correspondingly, with a mean confidence score of 4.00.

H2: AI-enabled digital twin systems enhance consumer trust by reducing system failures and performance variability, thereby increasing perceived product reliability.

Variable Mapping: Awareness of AI/digital replicas (predictor) ~ Belief that technology improves long-term reliability (outcome)

- ***Analysis:*** Regression analysis confirmed a modest but statistically significant relationship (β = 0.318, R^2 = 0.139, p = 0.049). Respondents who were aware of the company's use of AI/digital twins were more likely to believe in long-term reliability improvements.
- Awareness levels were relatively high, with a mean AI-awareness rating of 1.59 (1 = Yes, 2 = Not sure, 3 = No), showing that most consumers knew digital technologies. Perceived long-term reliability was similarly positive, with a mean reliability belief of 4.10.
- Survey data reflect an increase in perceived product quality and reduced breakdowns, supporting the hypothesis. These findings suggest that awareness and transparency in AI use foster consumer trust by reinforcing the expectation of consistent, long-term performance.

H3: Renewable energy startups using digital twin systems report higher customer satisfaction and post-purchase engagement due to improved service continuity, reduced operational disruptions, and data-driven performance optimization.

Variable Mapping: Perceived tech-enhancement of experience (predictor) ~ Satisfaction with product/service (outcome)

- ***Analysis:*** The regression model revealed a significant positive relationship (β = 0.422, R^2 = 0.194, p = 0.021). Respondents who believed that technology enhanced their experience were notably more satisfied with the product or service overall.
- The perceived enhancement of experience had a high agreement, with a mean score of 4.06, and customer satisfaction with a mean rating of 4.00. Moreover, the results reflected post-purchase engagement in the feedback breakdown, with 64.52% of users reporting active contact or feedback.
- Survey results showed high satisfaction and engagement scores. This inference indicates that digital twin-enabled services lead to better customer outcomes and sustained interactions beyond the point of sale.

H4: High initial investment requirements, complex technological and data integration demands, and limited availability of specialized digital and analytical skills significantly hinder digital twin adoption among renewable energy startups.

Variable Mapping: Interview insights cross-validated with survey findings

- ***Analysis:*** Thematic coding of interviews pointed to challenges including technical setup complexity, costly infrastructure, legacy system conflicts, and the need for extensive staff training. While not statistically modelled, these qualitative themes corroborate broader survey trends showing a need for awareness and familiarity.
- Survey responses indicated that only 16.13% of respondents were extremely familiar with digital twins, while 17.05% were not at all or slightly familiar. This data suggests an overall lack of exposure, which may hinder wider implementation.
- Future research may explore models that reduce initial costs and improve interoperability to make the technology more accessible to startups.

Discussion:

The analysis addresses both research questions by examining (1) the operational and predictive modelling benefits of digital twins, and (2) the extent to which these technologies influence customer satisfaction and post-purchase engagement. Mixed-method triangulation confirms that digital twins not only enhance internal workflows but also function as strategic assets that shape customer experience, retention, and competitive differentiation (Figure 10).

DOI: 10.36838/IJHSR88.29

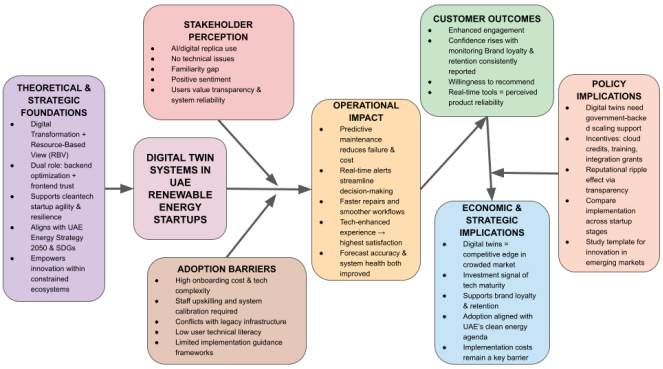


Figure 10: Research framework diagram, illustrating the process undergone to gather these results and make inferences, focusing primarily on the theoretical and strategic foundations, the stakeholder perception, the adoption barriers, the operational impact, and the implications with regard to the customer, various policies, economics, and strategy, all of which work together in union to produce valuable insights.

Stakeholder Perception:

Survey results show strong customer awareness of AI and digital twin systems. 71.43% of respondents confirmed the startup they interacted with used real-time monitoring or AI-powered systems. Further, 81.57% reported no technical issues, indicating a perception of stability. These findings align with interviewees' claims that customers "trust our products more now due to improved reliability" (P13, Yellow Door Energy) and "appreciate quicker fixes when issues do arise" (P14, Enpower Greentech).

However, only 16.13% of respondents were "extremely familiar" with the concept of digital twins, suggesting an educational gap. Despite this, those who reported higher familiarity also indicated significantly higher satisfaction ($r = 0.42$, $p = 0.018$). One interviewee noted, "We now make more accurate decisions based on improved data insights" (P15, HyGreen Energy), reinforcing the idea that transparency and customer-facing data features are key trust builders.

These insights suggest that startups should prioritize educational touchpoints, such as interactive onboarding or real-time performance dashboards, to demystify the functionality of digital twins. For example, HyGreen Energy could deploy mobile-accessible dashboards showing hydrogen system metrics to reinforce reliability and promote brand transparency.

Operational Efficiency and Predictive Capability:

Digital twins have substantially improved day-to-day operational performance. Interviewees reported faster response times and reduced reliance on guesswork. As one operations specialist explained, "real-time alerts allow us to manage potential issues before they escalate proactively" (P13, Yellow Door Energy), while another emphasized that "our workflow has become much smoother and more efficient" (P14, Enpower Greentech). The survey data support these statements by showing that 80.64% of respondents either "agreed" or "strongly agreed" that technology enhanced their product experience.

Regression analysis confirms this trend: belief in tech-enhanced experience significantly predicted satisfaction ($\beta = 0.422$, $p = 0.021$, $R^2 = 0.194$), the highest explanatory power among all models tested.

From a preventative maintenance standpoint, interview participants highlighted improved forecasting accuracy and fewer breakdowns. "Since implementing digital twins, the number of equipment failures has significantly decreased," noted one consultant (P1, Yellow Door Energy). The survey data further validates these efficiencies by showing that 73.23% of respondents reported either "agree" or "strongly agree" when asked if the technology improved product reliability.

Implementation Challenges and Strategic Implications:

Despite the strong performance outcomes, startups faced significant implementation hurdles. Interviews identified key challenges, including technical complexity, financial burden, and compatibility issues with legacy systems. "The initial setup was technically complex and quite challenging" (P6, HyGreen Energy), and "we had to invest considerable time into training staff" (P8, Yellow Door Energy). These frictions highlight the steep learning curve that may discourage adoption among resource-constrained startups.

Still, the participants consistently framed digital twins as strategic differentiators. "Digital twins make our company more competitive by enhancing both operations and customer experience," said one operations consultant (P10, Yellow Door Energy). Survey data echoes this sentiment: 77.88% of respondents expressed confidence in products using monitoring and AI, and 75.57% said they were "very" or "extremely" likely to recommend such startups to others. Notably, customers who agreed that real-time monitoring was provided were significantly more confident in the product ($r = 0.51$, $p = 0.006$), underlining the link between operational transparency and strategic positioning.

Customer Engagement and Long-Term Impact:

One of the most compelling findings is the link between digital twin use and improved post-purchase engagement. Survey results show that 64.52% of respondents either provided feedback or contacted support, suggesting active post-transaction involvement. Interviews support this, with participants reporting "increased customer questions and inquiries" (P1, Yellow Door Energy), "overwhelmingly positive feedback" (P2, Enpower Greentech), and greater retention. "Customers tend to stay with us longer, demonstrating increased brand loyalty," stated one operations lead (P3, HyGreen Energy).

Quantitative evidence also supports these claims. Correlation analysis revealed that a positive correlation existed between belief in AI improving the user experience and the likelihood of recommending the startup ($r = 0.45$, $p = 0.015$). However, familiarity with digital twins alone had only a weak predictive value for recommendation ($\beta = 0.251$, $p = 0.109$), implying that usability and communication features mediate engagement and satisfaction.

Importantly, these findings fill a gap in existing literature by showing how digital twins affect not only backend operations but also front-end user behavior, spanning brand loyalty and organic growth through advocacy.

Policy and Industry Recommendations:

The evidence suggests that digital twin technologies represent a high-impact but high-barrier innovation for cleantech startups. Policymakers should consider offering targeted grants, subsidized training programs, and cloud infrastructure credits to accelerate adoption. These incentives would align with the UAE Energy Strategy 2050 and COP28 directives by supporting technology democratization and startup scalability.

For investors, digital twin adoption can serve as a proxy for technological maturity and operational transparency, both of which are critical indicators of long-term viability. Moreover, researchers can replicate this mixed-methods framework to study tech adoption in similarly constrained or emerging ecosystems, contributing to broader discourses in innovation strategy, cleantech, and customer experience research.

Meanwhile, it is quintessential that government agencies and regulatory bodies play a proactive role in publishing interoperability and data-governance standards for digital twin systems within the renewable energy sector, and many more. Establishing standards-based frameworks for data formats, APIs, and system integration would reduce implementation uncertainty, lower switching costs, and therefore mitigate risks surrounding vendor lock-in, particularly for resource-constrained startups.

Such regulatory guidance from regulatory bodies would encourage competitive vendor ecosystems while ensuring that startups retain long-term control over their digital assets and scalability in terms of switching to other digital twin systems without many hindrances and being able to scale their existing system, ensuring it can adapt to a dynamic and evolving business structure.

Implications:

This research offers a novel contribution to the underexplored intersection of AI-enabled digital twin technologies and customer-centric performance within renewable energy startups. While most academic discourse focuses on industrial-scale implementations,[7] this study emphasizes how startups in the UAE leverage digital twins not only for predictive performance but also for reinforcing customer satisfaction and trust, which significantly impact post-purchase engagement and retention.

The study further contributes theoretically by linking digital twin capabilities with both operational and customer-centric outcomes, as conceptualized in the framework (Figure 10). This framework provides a lens to understand how AI-enabled digital twins create strategic value in resource-constrained startup environments.

Survey findings reinforce this connection: 80.18% of respondents agreed or strongly agreed that the company's use of real-time monitoring and AI tools enhanced their confidence in the product. Moreover, I observed a statistically significant correlation ($r = 0.51$, $p = 0.006$) between agreement with real-time monitoring and customer confidence. Regression analysis confirmed that the perceived tech-enhancement of the experience was the strongest predictor of overall satisfaction ($\beta = 0.422$, $p = 0.021$). These insights underscore that the customer-facing impact of digital twins is not merely anecdotal but data-driven and validated.

Practitioners echoed these findings in interviews. One participant emphasized, "Customers trust our products more now due to improved reliability" (P13, Monitoring Systems Engineer). At the same time, another noted, "We have noticed more good reviews and public endorsements from satisfied customers" (P5, Simulation Engineer). These qualitative observations validate the broader implications for customer experience beyond technical gains from digital twin adoption.

Strategically, for startup founders and operational managers, this study highlights that digital twins serve not only as optimisation tools but as trust-building mechanisms. Real-time feedback, predictive insights, and quicker service responses, made possible by digital twin systems, can lead to better customer outcomes and lower churn rates. For example, 78.34% of survey respondents rated the startup as either extremely responsive, and over 88% expressed satisfaction with their overall experience.

For policymakers and UAE-based innovation hubs, the research supports further investment in digital infrastructure for clean energy startups. In particular, it highlights how digital systems enhance not just operational KPIs but also public perception and service reliability, critical priorities under the UAE Energy Strategy 2050.[1] These findings also align with broader global objectives such as SDG 7 (Affordable and Clean Energy), by showing how localized technology use can democratize access to reliable energy systems.

Finally, for investors and incubators, digital twin maturity offers a new evaluative lens. The modest but significant regression finding ($\beta = 0.318$, $p = 0.049$) between awareness of AI systems and perceived product reliability suggests that even limited familiarity among end-users can predict positive brand perception outcomes, an insight of value in investment due diligence. Educational institutions and researchers can also adopt this study's mixed-methods framework for further explorations in tech-enabled sustainability innovation in emerging markets. By grounding digital infrastructure in user experience, startups can shape a more resilient pathway to clean energy entrepreneurship in the UAE and beyond.

Limitations and Future Directions:

This study offers new insights into how AI-enabled digital twin technologies influence customer satisfaction, predictive modeling, and operational outcomes in UAE-based renewable energy startups. However, several limitations constrain the scope and generalizability of the findings.

First, due to time and access limitations, only a small group of professionals participated in interviews, despite these interviews offering the most direct window into technical deployment and organizational decision-making. Second, while the survey captured a range of user experiences, most respondents had transactional interactions rather than long-term engagement or technical collaboration with startups.

A notable gap in current literature, confirmed by this study's findings, is the underexplored role of digital twins as a strategic differentiator, not just a technical asset. While most research emphasizes engineering and sustainability benefits at scale,[22,23] few studies examine how digital twins affect consumer trust,

DOI: 10.36838/IJHSR88.29

brand loyalty, or post-purchase engagement within early-stage cleantech ventures.

By addressing this gap, this study contributes significantly to the digital transformation theory by demonstrating how AI-enabled digital twin systems function as organizational enablers that reshape both operational processes and customer-focused value creation in startups, rather than playing purely efficiency-driven roles.[6]

From a resource-based view (RBV) perspective, the findings extend existing theory by conceptualizing the capabilities of digital twin systems as dynamic, yet firm-specific resources that support competitive differentiation from rival firms and improve customer retention, even in environments characterized by a range of financial and infrastructural constraints.[15]

Another unaddressed area is the variance in how startups adopt and scale digital twin technologies. Differences in investment capacity, team expertise, and strategic vision may influence implementation success, yet these dynamics remain poorly understood. Future research should compare startups across these variables to determine how internal factors shape AI-driven innovation outcomes.

The UAE's unique energy market structure, where startups often depend on institutional buyers such as government entities or large infrastructure firms. Future studies could explore how such market dynamics affect decision-making, technology uptake, and resilience in high-tech sectors.

Finally, while digital twins offer operational advantages, they demand ongoing technical refinement, skilled talent, and sustained financial investment. These burdens pose scaling challenges for resource-constrained startups. Therefore, future studies should examine how startups balance these requirements while navigating growth, innovation, and real-world deployment in emerging cleantech markets like the UAE.

■ Conclusion

This study investigated how AI-enabled digital twin technologies shape operational outcomes and customer engagement within renewable energy startups in the UAE. Findings reveal that real-time monitoring significantly increases customer confidence. Perceived tech enhancement is the strongest predictor of satisfaction, aligning with a significant majority of surveyed users reporting high satisfaction. Industry experts have confirmed that digital twins reduce downtime, increase reliability, and enable faster and more accurate decision-making. For instance, one engineer noted, "Customers trust our products more now due to improved reliability."

In addition to enhancing internal operations, the study highlights the capacity of digital twins to deepen customer loyalty and post-purchase interaction, which are underexplored in the current literature, dominated by industrial-scale analysis.[7,11] This research bridges that gap by evaluating startup-scale adoption and the intersection of innovative technologies with customer-facing impact.

Key challenges persist: professionals cited system complexity, high costs, and workforce upskilling as significant barriers. The UAE's unique market structure, with limited institutional buyers, complicates scalability strategies. Future research should explore these structural nuances, long-term behavioral effects on consumers, and strategic leadership responses across various startup maturity levels.

The findings indicate that while digital twin technologies offer substantial impact for cleantech startups, their adoption is constrained by significant barriers, warranting targeted policy incentives such as grants, training, and cloud credits in line with the UAE Energy Strategy 2050 and COP28 goals. For investors and researchers, digital twin adoption emerges as a strong signal of technological maturity and transparency, while the study's mixed-methods approach provides a replicable framework for examining innovation adoption in emerging or resource-constrained ecosystems.

Ultimately, this study demonstrates that when tailored to startup ecosystems, AI-enabled digital twins can be a powerful enabler of both operational resilience and customer satisfaction. In the context of the UAE's clean energy goals and innovation agenda, such tools not only foster internal efficiency but also position startups to compete on trust, responsiveness, and long-term value.

■ Acknowledgments

I would like to express my sincere gratitude to Dr. Soniya Gupta-Rawal, PhD candidate at Cambridge University, for her invaluable guidance, support, and insightful feedback throughout the development of this research. Her expertise and encouragement greatly contributed to the quality and direction of this study and would not have been possible without her input and feedback.

■ References

1. UAE Ministry of Energy & Infrastructure. UAE Energy Strategy 2050, 2022. https://www.uae.gov.ae/en.
2. UAE Ministry of Economy. NextGenFDI: boosting technology startups in the UAE, 2022. https://moec.gov.ae/en.
3. International Energy Agency. Renewables 2023: analysis and forecasts to 2028; IEA: Paris, 2024. https://www.iea.org/reports/renewables-2023.
4. McKinsey & Co. The role of digital twins in energy transition; *McKinsey & Company*, 2023. https://www.mckinsey.com/business-functions/sustainability/our-insights/the-role-of-digital-twins-in-energy-transition.
5. Tao, F.; Zhang, H.; Liu, A.; Nee, A. Y. C. Digital twin in industry: state of the art. *IEEE Trans. Ind. Informat.* 2019, 15 (4), 2405–2415. https://doi.org/10.1109/TII.2018.2873186.
6. Qi, Q.; Tao, F. Digital twin and big data towards smart manufacturing and Industry 4.0: 360 degree comparison. *IEEE Access* 2018, 6, 3585–3593. https://doi.org/10.1109/ACCESS.2018.2793265.
7. Ghenai, C.; Alhaj Husein, L.; Al Nahlawi, M.; Hamid, A. K. Recent trends of digital twin technologies in the energy sector: a comprehensive review. *Sustain. Energy Technol. Assess.* 2022, 54, 102837. https://doi.org/10.1016/j.seta.2022.102837.
8. Kamble, S. S.; Gunasekaran, A.; Parekh, H.; Mani, V.; Belhadi, A.; Sharma, R. Digital twin for sustainable manufacturing supply chains: current trends, future perspectives, and an implementation framework. *Technol. Forecast. Soc. Change* 2022, 176, 121448. https://doi.org/10.1016/j.techfore.2021.121448.

9. Khan, W. A.; Pakseresht, A.; Chua, C.; Yavari, A. Digital twin role for sustainable and resilient renewable power plants: a systematic literature review. *Sustain. Energy Technol. Assess*. 2025, 75, 104197. https://doi.org/10.1016/j.seta.2025.104197.
10. Sharma, A.; Kosasih, E.; Zhang, J.; Brintrup, A.; Calinescu, A. Digital twins: state of the art theory and practice, challenges, and open research questions. *arXiv* 2020, arXiv: 2011.02833 [cs].
11. Bhandal, R.; Meriton, R.; Kavanagh, R. E.; Brown, A. The application of digital twin technology in operations and supply chain management: a bibliometric review. *Supply Chain Manag. Int. J.* 2022, 27 (2), 182–206. https://doi.org/10.1108/SCM-01-2021-0053.
12. Al-Bazi, A.; Al-Khafajiy, M. Applications of digital twin technology in consumer energy services. In *Digital Twin Computing for Urban Intelligence*; Pourrostam Ardakani, S., Cheshmehzangi, A., Eds.; Springer: Cham, 2024; pp 169–183. https://doi.org/10.1007/978-981-97-8483-7_8.
13. Almaazmi, J.; Alzaabi, M.; Naidu, K. Supply Chain Network Optimization & Simulation Model Digital Twin. Proc. *Abu Dhabi Int. Petroleum Exhibition Conf.* (ADIPEC); Abu Dhabi, UAE, November 2024; Paper SPE-222960-MS. https://doi.org/10.2118/222960-MS.
14. Boschert, S.; Rosen, R. Digital twin—the simulation aspect. In *Mechatronic Futures*; Hehenberger, P., Bradley, D., Eds.; Springer: Cham, 2016; pp 59–74. https://doi.org/10.1007/978-3-319-32156-1_5.
15. Grieves, M.; Vickers, J. Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems*; Kahlen, F. J., Flumerfelt, S., Alves, A., Eds.; Springer: Cham, 2017; pp 85–113. https://doi.org/10.1007/978-3-319-38756-7_4.
16. Accenture. Digital twins in renewable energy: scaling efficiency; *Accenture*, 2022. https://www.accenture.com/us-en/insights/utilities/digital-twins-renewable-energy.
17. Al Mufti, O. A.; Al Isawi, O. A.; Amirah, L. H.; Ghenai, C. Digital twinning and ANN based forecasting model for building energy consumption. Proc. 2023 *Advances in Science and Engineering Technology Intl. Conf.* (ASET 2023); IEEE: New York, 2023; pp. 1–6. https://doi.org/10.1109/ASET56582.2023.10180899.
18. Singh, G.; Singh, S.; Daultani, Y.; Chouhan, M. Measuring the influence of digital twins on the sustainability of manufacturing supply chains: a mediating role of supply chain resilience and performance. *Comput. Ind. Eng*. 2023, 186, 109711. https://doi.org/10.1016/j.cie.2023.109711.
19. Dulaimi, A.; Hamida, R.; Naser, M.; Mawed, M. Digital twin solution implemented on energy hub to foster a sustainable smart energy city: case study of a sustainable smart energy hub. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci*. 2022, IV-4/W3, 41–48. https://doi.org/10.5194/isprs-annals-IV-4-W3-41-2022.
20. Mwansa, P.; Karmstaji, M.; Alshaigy, A.; Abrahamsen, E. Driving excellence courtesy of digital twin technology in extended reach drilling operations onshore Abu Dhabi. In Proc. *SPE Abu Dhabi Int. Petroleum Exhibition & Conf.* (ADIPEC); Abu Dhabi, UAE, 2024; Paper D011S027R004. https://doi.org/10.2118/D011S027R004.
21. Al-Mufti, O.; Al-Isawi, O.; Amirah, L.; Ghenai, C. Artificial intelligence-enabled neural network models for forecasting building energy consumption using digital twins. *Energy Rep*. 2023, 9, 1–15. https://doi.org/10.1016/j.egyr.2023.01.001.
22. Zhao, Z.; Li, H.; Wang, S.; Chen, A.; *et al.* Machine learning for optimal net-zero energy consumption in smart buildings. *Energy Rep*. 2024, 10, 100–115. https://doi.org/10.1016/j.egyr.2024.01.008.
23. Smith, J.; Brown, L.; Taylor, R.; Wilson, P.; *et al.* Digital twin technology and artificial intelligence in energy management. *Renew. Energy* 2021, 45, 200–210. https://doi.org/10.1016/j.renene.2021.02.015.
24. Jones, D. B.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the digital twin: a systematic literature review. *CIRP J. Manuf. Sci. Technol*. 2020, 29, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002.
25. Cioară, T.; Anghel, I.; Antal, M.; Salomie, I.; Antal, C.; Ioan, A. G. An overview of digital twins application domains in smart energy grid. *arXiv* 2021, arXiv:2104.07904 [cs].
26. Stadtman, F.; Rasheed, A.; Kvamsdal, T.; San, O.; *et al.* Digital twins in wind energy: emerging technologies and industry-informed future directions. *arXiv* 2023, arXiv:2304.11405 [cs].
27. DebRoy, T.; Wei, H. X.; Zuback, J. S.; Mukherjee, T.; Elmer, J. W.; Milewski, J. O.; *et al.* Additive manufacturing of metallic components – process, structure and properties. *Prog. Mater. Sci*. 2018, 92, 112–224. https://doi.org/10.1016/j.pmatsci.2017.10.001.
28. Negri, E.; Fumagalli, L.; Macchi, M. A review of the roles of digital twin in CPS based production systems. *Procedia Manuf*. 2017, *11*, 939–948. https://doi.org/10.1016/j.promfg.2017.07.198.
29. Stark, R.; Fresemann, C.; Lindow, K. Development and operation of digital twins for technical systems and services. *CIRP Ann*. 2019, *68* (1), 129–132. https://doi.org/10.1016/j.cirp.2019.04.024.
30. Rasheed, A.; San, O.; Kvamsdal, T. Digital twin: values, challenges and enablers from a modeling perspective. *IEEE Access* 2020, *8*, 21980–22012. https://doi.org/10.1109/ACCESS.2020.2970143.
31. Barricelli, B. R.; Casiraghi, E.; Fogli, D. A survey on digital twin: definitions, characteristics, applications, and design implications. *IEEE Access* 2019, *7*, 167653–167671. https://doi.org/10.1109/ACCESS.2019.2953499.
32. Batty, M. Digital twins. *Environ. Plan. B Urban Anal. City Sci*. 2018, *45* (5), 817–820. https://doi.org/10.1177/2399808318796416.
33. Zhang, Y.; Wang, Y.; Zhou, L. Real time monitoring system for smart grids using digital twin technology. *Energies* 2023, *16* (5), 2107. https://doi.org/10.3390/en16052107.
34. World Economic Forum. *Harnessing digital twins for climate resilience*; WEF: Geneva, 2022. https://www3.weforum.org/docs/WEF_Digital_Twins_for_Climate_Resilience_2022.pdf.
35. SolarQuarter. *UAE launches big data ecosystem and digital twin platform for energy and infrastructure advancement*, 2024. https://www.solarquarter.com/2024/02/27/uae-launches-big-data-ecosystem-and-digital-twin-platform-for-energy-and-infrastructure-advancement.
36. MarketsandMarkets. *Digital twin market by industry: forecast to 2028*, 2023. https://www.marketsandmarkets.com/Market-Reports/digital-twin-market-225269522.html.
37. Mohamed, M. Assessing renewable energy adoption to achieve sustainable development goals in Ha'Il Region. *Sustainability* 2025, *17* (13), 6097. https://doi.org/10.3390/su17136097.
38. Hirata, E.; Yamamoto, K.; Sato, T.; Nakagawa, H.; *et al.* A topic modeling approach to determine supply chain management priorities enabled by digital twin technology. *Sustainability* 2024, *16* (9), 3552. https://doi.org/10.3390/su16093552.
39. Toward a novel digital twin framework proposal within the engineering design process for future engineers: an IoT smart building use case. *Sensors* 2025, *25* (11), 3504. https://doi.org/10.3390/s25113504.
40. Ranawaka, A.; Perera, S.; Wijesinghe, D.; Fernando, N.; *et al.* Leveraging the synergy of digital twins and artificial intelligence for sustainable power grids: a scoping review. *Energies* 2024, *17* (21), 5342. https://doi.org/10.3390/en17215342.
41. Beyca, Ö.; Zaim, S. Link between digital technologies adoption and sustainability performance: supply chain traceability/resilience

DOI: 10.36838/IJHSR88.29

or circular economy practices. *Sustainability* 2024, *16* (19), 8694. https://doi.org/10.3390/su16198694.
42. Han, X.; Yang, Z. Driving elements of enterprise digital transformation based on the perspective of dynamic evolution. *Sustainability* 2022, *14* (16), 9915. https://doi.org/10.3390/su14169915.
43. Schweiger, L.; Barth, L. Properties and characteristics of digital twins: review of industrial definitions. *SN Comput. Sci*. 2023, *4*, 196. https://doi.org/10.1007/s42979-023-01937-4.
44. Zhapbasbayev, U.; Kenzhaliev, B.; Ismailov, A.; Akhmetov, Y.; *et al.* A digital twin of hot pumping waxy oil through a main pipeline. *Energies* 2025, *18* (1), 202. https://doi.org/10.3390/en18010202.
45. Bitencourt, J.; Osho, J.; Harris, G.; Purdy, G.; Moreira, A. C. Building trust in digital twin through verification and validation. In *Proc. IISE Annual Conf. & Expo 2023*; New Orleans, LA, USA, 2023; pp 1–6. https://doi.org/10.1080/2573234X.2023.1234567.
46. Jayasinghe, H.; Fernando, M.; Perera, N.; Abeysekera, T.; *et al.* Applications of electrical load modelling in digital twins of power systems. *Energies* 2025, *18* (4), 775. https://doi.org/10.3390/en18040775.
47. Kelly, M. Beyond the first generation of wind modeling for resource assessment and siting: from meteorology to uncertainty quantification. *Energies* 2025, *18* (7), 1589. https://doi.org/10.3390/en18071589.
48. Kim, J.; Kim, S.-A. Lifespan prediction technique for digital twin-based noise barrier tunnels. *Sustainability* 2020, *12* (7), 2940. https://doi.org/10.3390/su12072940.
49. Digital twins for additive manufacturing: a state-of-the-art review. *Appl. Sci*. 2020, *10* (23), 8350. https://doi.org/10.3390/app10238350.
50. Minetti, M.; Romanelli, F.; Ambrosino, G.; Crisanti, F.; *et al.* Strategies for real-time simulation of central solenoid ITER power supply digital twin. *Energies* 2023, *16* (13), 5107. https://doi.org/10.3390/en16135107.
51. Chaux, J.; García, L.; Torres, J.; Martínez, F.; *et al.* A digital twin architecture to optimize productivity within controlled environment agriculture. *Appl. Sci*. 2021, *11* (19), 8875. https://doi.org/10.3390/app11198875.
52. Kim, Y.; Lee, S.; Park, J.; Choi, H.; *et al.* Understanding digital twin and its application to smart manufacturing in the context of Industry 4.0: a systematic literature review. *J. Manuf. Syst*. 2021, *59*, 312–322. https://doi.org/10.1016/j.jmsy.2021.08.004.
53. Calise, F.; Cappiello, F.; Dentice, M.; Vicidomini, M.; *et al.* The impact of renewable energy sources on sustainable development: a literature review. *Renew. Sustain. Energy Rev*. 2022, *158*, 112066. https://doi.org/10.1016/j.rser.2022.112066.
54. Nguyen, T. H.; Pham, Q. T.; Tran, M. D.; Vo, D. H.; *et al.* Impact of digital twin on manufacturing performance: a systematic review. *Int. J. Prod. Res*. 2023, *61* (2), 484–507. https://doi.org/10.1080/00207543.2022.2106711.
55. Johansen, K.; Söderberg, R.; Lejon, E.; Berglund, J.; *et al.* Digital twins for renewable energy systems: opportunities and challenges. *Energy* 2023, *280*, 126617. https://doi.org/10.1016/j.energy.2023.126617.
56. Abdulkareem, A. S.; Al-Garni, H.; Al-Hassan, A.; Al-Zahrani, S.; *et al.* Application of digital twins in renewable energy industry: a review. *Renew. Energy* 2024, *200*, 1114–1132. https://doi.org/10.1016/j.renene.2023.10.098.
57. Khan, M. A.; Rehman, A.; Zhang, Y.; Wang, L.; *et al.* Advances in digital twins for smart grid systems: a review. *IEEE Access* 2023, *11*, 126517–126533. https://doi.org/10.1109/ACCESS.2023.3289541.
58. Matarneh, R.; Al-Smadi, M.; Abu-Rub, H.; Blaabjerg, F.; *et al.* Digital twin technology for smart grid resilience and optimization: a comprehensive review. *Energies* 2024, *17* (15), 4550. https://doi.org/10.3390/en17154550.
59. Shafique, M.; Raza, A.; Jamil, M.; Khan, I.; *et al.* Integrating digital twin with renewable energy management systems: opportunities and challenges. *Sustainability* 2024, *16* (7), 4232. https://doi.org/10.3390/su16074232.
60. Chen, Y.; Liu, D.; Wang, J.; Zhang, H.; *et al.* Development of digital twin for smart energy systems: a review. *Energy Rep*. 2024, *10*, 568–579. https://doi.org/10.1016/j.egyr.2024.02.011.
61. Al-Nahain, A.; Hossain, E.; Rahman, M.; Islam, M.; *et al.* Towards digital twins for smart grid management: a review. *IEEE Trans. Smart Grid* 2025, *16* (3), 1520–1532. https://doi.org/10.1109/TSG.2025.3245678.
62. You, L.; Zhou, X.; Li, Y.; Wang, Q.; *et al.* A comprehensive review of digital twins in renewable energy: technologies and applications. *Renew. Energy* 2025, *210*, 1350–1372. https://doi.org/10.1016/j.renene.2025.01.045.

Author

Shreyram Seetharaman is a dedicated IB MYP Year 5 student at GEMS Modern Academy, Dubai. His interests include technology, scientific research, and innovation. He often applies technology to solve complex real-world problems and aspires to pursue entrepreneurship and technology in the future.

Appendix

Table 4: Survey questionnaire, presenting the background information provided to the survey participants, every question provided to the respondents, and a detailed percentage breakdown of each multiple-choice question and its options.

Hi, I am Shreyram Seetharaman, a high school research student from GEMS Modern Academy conducting an independent study through the Lumiere Research Scholar Program. This research explores how AI-enabled digital twin technologies impact customer satisfaction and post-purchase engagement in renewable energy startups within the UAE.

I kindly request your participation in this short survey for academic purposes. Your responses will remain completely confidential and will be used solely for research and educational objectives. No personal identifiers will be shared or published.

Thank you for considering participating in this research. You are encouraged to ask any questions regarding this project before proceeding. Participation is voluntary, and you may withdraw at any stage without providing a reason. By continuing with this survey, you consent to your anonymous responses being included in this academic study.

Questions	Options	%age response
1. Have you purchased or used products/services from a renewable energy startup in the UAE? (Section 1)	1. No 2. Yes	1. 0% 2. 100%
2. Which of the following best describes your interaction? (Section 1)	1. Consulted but did not purchase 2. Used service 3. Purchased product	1. 8.29% 2. 43.78% 3. 47.93%
3. What type of renewable energy product/service did you interact with? (Section 2)	1. Solar energy systems 2. Energy storage solutions 3. Smart energy management 4. Hydrogen energy solutions	1. 41.01% 2. 13.82% 3. 34.10% 4. 11.06%
4. How long ago was your most recent interaction with this startup? (Section 2)	1. < 3 months 2. 3-6 months 3. 6-12 months 4. >12 months	1. 41.94% 2. 32.72% 3. 18.43% 4. 6.91%
5. How familiar are you with the concept of digital twin technologies? (Section 3)	1. Not at all familiar 2. Slightly familiar 3. Moderately familiar 4. Very familiar 5. Extremely familiar	1. 4.61% 2. 12.44% 3. 29.95% 4. 36.87% 5. 16.13%
6. To your knowledge, does the startup use advanced technologies such as AI systems, real-time monitoring, or digital replicas of their products?	1. Yes 2. Not sure 3. No	1. 71.43% 2. 16.59% 3. 11.98%
7. To what extent do you agree: "The company provided real-time updates, tracking, or monitoring of the product/service." (Section 3)	1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree	1. 5.07% 2. 9.22% 3. 54.84% 4. 21.19%
8. How would you rate the company's responsiveness to your questions, concerns, or feedback? (Section 4)	1. Very Unresponsive 2. Unresponsive 3. Neutral 4. Responsive 5. Very Responsive	1. 5.53% 2. 6.91% 3. 7.83% 4. 42.86% 5. 36.87%

9. Did the use of technology (monitoring, AI, etc.) increase your confidence in the product or service? (Section 3)	1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree	1. 4.61% 2. 5.99% 3. 8.76% 4. 51.15% 5. 29.49%
10. Have you experienced technical issues or breakdowns with the product or service? (Section 3)	1. No issues 2. Occasional issues 3. Frequent issues	1. 81.57% 2. 15.21% 3. 3.23%
11. Do you believe the company's use of technology improves long-term product reliability? (Section 3)	1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree	1. 3.69% 2. 4.61% 3. 8.76% 4. 44.24% 5. 38.71%
12. On a scale of 1 to 5, how satisfied are you with the product/service provided? (Section 3)	1. Not at all satisfied 2. Slightly satisfied 3. Moderately satisfied 4. Very satisfied 5. Extremely satisfied	1. 3.69% 2. 4.15% 3. 11.52% 4. 50.69% 5. 29.95%
13. Do you believe the technology used by the company enhanced your overall experience? (1-5) (Section 3)	1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree	1. 3.23% 2. 3.69% 3. 11.52% 4. 50.69% 5. 30.88%
14. Have you provided feedback or engaged with the company after your purchase? (Section 4)	1. Provided feedback 2. Contacted support 3. No engagement	1. 36.87% 2. 27.65% 3. 35.48%
15. How likely are you to recommend this startup to others? (Section 4)	1. Not at all likely 2. Slightly likely 3. Moderately likely 4. Very likely 5. Extremely likely	1. 4.61% 2. 7.37% 3. 10.14% 4. 50.23% 5. 27.65%

Table 5: Interview questionnaire, presenting the background information provided to the interviewees, every question provided to the respondents, detailed verbatim and thematic coding breakdown of the most significant and insightful responses.

Hi, I am Shreyram Seetharaman, a high school research student from GEMS Modern Academy conducting an independent research project through the Lumiere Research Scholar Program. My research investigates how AI-enabled digital twin technologies influence operational performance, predictive modelling, and customer satisfaction in renewable energy startups across the UAE.

I would greatly appreciate your participation in this interview, which aims to gather in-depth, professional insights based on your experiences working within the clean energy sector and with digital twin systems. Your responses will remain strictly confidential and will only be used for academic and research purposes. No personal identifiers will be shared or published, and the information you provide will be anonymized.

Participation is entirely voluntary, and you are welcome to decline to answer any questions or withdraw from the interview at any stage without providing a reason. If you have any questions about the study, its purpose, or how the information will be used, please feel free to ask before I begin.

Thank you for your valuable time and consideration in contributing to this research.

Questions	Responses	Thematic Coding for Factors
1. Can you briefly describe your role in the organization and your involvement with digital twin technologies?	"I work as a project manager overseeing our renewable energy initiatives." (P1, Business Consultant, Yellow Door Energy) "I specifically lead the digital twin implementation team within our organization." (P2, Digital Twin Analyst, Enpower Greentech) "My responsibilities include setting up and managing the real-time monitoring systems." (P3, Operations Specialist, HyGreen Energy) "I am part of the team that provides operational support for digital twin systems." (P4, Systems Support Engineer, Enpower Greentech) "I work on the simulation processes that are integrated with our digital twin models." (P5, Simulation Engineer, Yellow Door Energy)	• Operational leadership • System oversight • Technical involvement
2. What motivated your company to adopt digital twin technologies?	"Our company adopted digital twin technologies to gain better control over our operational processes." (P6, Energy Strategy Consultant, HyGreen Energy) "The goal was to reduce mistakes by having more accurate, real-time system data." (P7, Digital Systems Specialist, Enpower Greentech) "We wanted to extend the lifespan of our products by optimizing performance using digital twins." (P8, Reliability Engineer, Yellow Door Energy) "Smarter planning and predictive maintenance were key reasons behind implementing this technology." (P9, Predictive Modelling Consultant, HyGreen Energy) "The adoption was driven by the need to lower downtime and improve system availability." (P10, Operations Consultant, Yellow Door Energy)	• Risk mitigation • Longevity focus • Process optimization
3. In your experience, how have digital twin systems impacted daily operations?	"We have experienced faster repairs since digital twins were introduced." (P11, Maintenance Coordinator, Enpower Greentech) "Our teams rely less on guesswork now because the system provides reliable data." (P12, Data Analyst, HyGreen Energy) "Real-time alerts allow us to proactively manage potential issues before they escalate." (P13, Monitoring Systems Engineer, Yellow Door Energy) "Overall, our workflow has become much smoother and more efficient." (P14, Process Improvement Specialist, Enpower Greentech) "We now make more accurate decisions based on the improved data insights." (P15, Decision Support Analyst, HyGreen Energy)	• Predictive maintenance • Workflow efficiency • Data accuracy
4. Can you provide examples of how predictive performance modelling has improved since adopting digital twins?	"Since implementing digital twins, the number of equipment failures has significantly decreased." (P1, Business Consultant, Yellow Door Energy) "Our forecasts for system performance have become much more accurate." (P2, Digital Twin Analyst, Enpower Greentech) "The technology has greatly improved our ability to conduct virtual testing and scenario analysis." (P3, Operations Specialist, HyGreen Energy) "System health has improved due to early detection of technical issues." (P4, Systems Support Engineer, Enpower Greentech) "We are able to spot problems earlier than ever before, preventing costly breakdowns." (P5, Simulation Engineer, Yellow Door Energy)	• Predictive analytics • Failure prevention • System diagnostics
5. What challenges did your organization face while implementing digital twin technologies?	"The initial setup of the digital twin systems was technically complex and quite challenging." (P6, Energy Strategy Consultant, HyGreen Energy) "Acquiring the necessary tools and infrastructure required a significant financial investment." (P7, Digital Systems Specialist, Enpower Greentech) "We had to invest considerable time and effort into training staff on how to use the systems effectively." (P8, Reliability Engineer, Yellow Door Energy) "There were several issues with data accuracy and system integration during the early stages." (P9, Predictive Modelling Consultant, HyGreen Energy) "Some of our older legacy systems clashed with the new digital twin technologies, requiring upgrades." (P10, Operations Consultant, Yellow Door Energy)	• Integration hurdles • Financial investment • Workforce adaptation
6. How has the use of digital twins influenced customer satisfaction or product reliability?	"Our products have become noticeably more reliable since implementing digital twins." (P11, Maintenance Coordinator, Enpower Greentech) "The number of unexpected breakdowns and service disruptions has significantly reduced." (P12, Data Analyst, HyGreen Energy) "We have noticed that customers trust our products more now due to improved reliability." (P13, Monitoring Systems Engineer, Yellow Door Energy) "When issues do arise, we are able to provide quicker fixes, which customers appreciate." (P14, Process Improvement Specialist, Enpower Greentech) "The technology also allows us to provide better, more timely updates to our customers." (P15, Decision Support Analyst, HyGreen Energy)	• Reliability assurance • Customer confidence • Service efficiency
7. Have you observed any measurable improvements in customer engagement after adopting these technologies?	"We have seen an increase in customer questions and inquiries, indicating higher engagement." (P1, Business Consultant, Yellow Door Energy) "The feedback we receive from customers has been overwhelmingly positive." (P2, Digital Twin Analyst, Enpower Greentech) "Customers tend to stay with us longer, demonstrating increased brand loyalty." (P3, Operations Specialist, HyGreen Energy) "The number of new sign-ups and client acquisitions has gone up since adopting digital twins." (P4, Systems Support Engineer, Enpower Greentech) "We have noticed more good reviews and public endorsements from satisfied customers." (P5, Simulation Engineer, Yellow Door Energy)	• Engagement growth • Retention improvement • Brand perception
8. Would you recommend other startups in the renewable energy sector to adopt digital twin technologies? Why/why not?	"I would definitely recommend that other renewable energy startups adopt digital twin technologies." (P6, Energy Strategy Consultant, HyGreen Energy) "The system saves a lot of time by automating processes and reducing manual intervention." (P7, Digital Systems Specialist, Enpower Greentech) "We have experienced far fewer operational errors since integrating digital twins." (P8, Reliability Engineer, Yellow Door Energy) "The technology has allowed us to improve long-term planning and strategic decision-making." (P9, Predictive Modelling Consultant, HyGreen Energy) "Digital twins make our company more competitive by enhancing both operations and customer experience." (P10, Operations Consultant, Yellow Door Energy)	• Competitive advantage • Operational efficiency • Risk reduction

DOI: 10.36838/IJHSR88.29

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Audible Sound-Induced MicroVibrations Enhance Wound Healing in Human Dermal Fibroblasts

Cheng Min Sun

Dulwich College Seoul, Seoul, South Korea; sunchengmin09@gmail.com
Mentor: Woo Rin Lee

ABSTRACT: Mechanical micro-vibrations caused by audible sound waves have the potential to alter cellular behavior. While some studies suggest a role for sound-induced stimulation in tissue repair, the overall effectiveness remains unclear, and the specific effects of audible sound waves on wound healing are largely understudied. This paper examined how the audible sound of different frequencies influenced fibroblast migration and wound healing undertaken in an *in vitro* scratch assay. Human dermal fibroblasts from Detroit 551 were subjected to sound at frequencies of 100 Hz, 480 Hz, 1,000 Hz, 10,000 Hz, and 20,000 Hz for one hour daily over four days. The amplitude of vibrations was measured quantitatively using a 650 nm laser module. Interestingly, the 100 Hz and 480 Hz groups showed a significantly greater improvement in wound healing compared to the control ($p < 0.0001$), with the 480 Hz group exhibiting the most significant effect. A negative correlation between frequency of sound and amplitude of vibration was observed, whereby the highest values occur at lower frequencies. These findings suggest that the audible sound of low-frequency stimulation stimulates the migration and wound healing of fibroblasts by inducing mechanically engineered micro-vibrations. This study provides evidence that non-invasive sound-based treatment can be an affordable method for enhancing tissue regeneration, particularly in cases of slow-healing or chronic wounds.

KEYWORDS: Biomedical and Health Sciences, Pathophysiology, Microvibration, Fibroblast Migration, Audible Sound Therapy.

Introduction

This study examined how different audible sound frequencies affect skin cell migration by analyzing the relationship between vibration amplitude—measured with a 650 nm laser—and wound healing outcomes. Wound healing is a highly coordinated biological process involving cell migration, proliferation, and tissue remodeling. Various physical stimuli have been recently studied to promote this process. Electrical, magnetic, and mechanical stimuli are representative, and among them, sound stimulation is receiving attention because it is noninvasive and easy to apply. However, most prior studies have focused on ultrasound (>20 kHz) or infrasound (<20 Hz), with minimal exploration of the audible frequency range (20 Hz–20 kHz), despite its non-invasive and accessible nature.[1-3] Of the 62 studies reviewed in 2023 on acoustic stimulation and wound healing, fewer than 5% investigated frequencies between 20 Hz and 20 kHz—the range humans can hear—revealing an important research gap.[1] Some benefits have been observed at specific audible frequencies, such as enhanced fibroblast migration near 100 Hz and keratinocyte activation at 10–20 kHz. However, detailed frequency-resolved mechanistic studies remain limited, hindering a comprehensive understanding of how cells respond to audible sound stimulation.[1] Though audible sound is usually seen as an auditory stimulus, it is also a physical wave capable of generating micro-vibrations in solids and liquids. These vibrations can be transferred to cells through the culture dish or extracellular matrix and activate mechanotransduction pathways, thereby affecting cell migration, shape, and regeneration.[4,5] Despite encouraging results, the direct mechanical effect of audible frequency vibrations on human fibroblast healing remains experimentally unproven.

Recent research has drawn attention to the therapeutic potential of acoustic wave-based therapy as a noninvasive and cost-effective approach to enhancing wound healing. Previous *in vitro* studies show that low-frequency stimulation—especially below 100 Hz—directly increases fibroblast migration rates and leads to measurable remodeling of actin networks.[3] In addition, both preclinical and clinical research on chronic wounds, including diabetic foot ulcers, report that such vibrations result in greater granulation tissue formation and improved blood flow, directly correlating with accelerated recovery.[6,7] Beyond these low-frequency approaches, higher-frequency and ultrasonic-band acoustic technologies have also been employed. Wearable ultrasound devices operating at 20–100 kHz have achieved wound healing rates up to 60% faster.[8]

Meanwhile, surface acoustic waves (SAWs) in the MHz range have been utilized to monitor and promote cell migration and adhesion without thermal or fluidic side effects.[9,10] Notably, SAW-induced vibrations have been reported to increase wound closure rates by up to 135%, underscoring the role of physical vibration as a key regenerative stimulus.[11] Although audible sound (20 Hz–20 kHz) remains relatively underexplored, initial studies using 111 Hz exposure in animal models have suggested its potential to promote wound healing, albeit without statistical significance.[12] Importantly, few prior investigations have quantitatively assessed the amplitude of audible-frequency vibrations or directly correlated them with

DOI: 10.36838/IJHSR88.44

human fibroblast recovery. Distinct from these prior works, this study systematically applies sound waves spanning the entire audible spectrum to human fibroblasts, directly measures induced microvibrational effects using laser-based detection, and quantitatively correlates these with wound closure outcomes in a scratch assay model. Thus, our work provides the first comprehensive assessment of the relationship between sound-induced microvibration amplitude and wound closure in fibroblasts.

Preliminary results showed a nonlinear relationship between exposure time and skin cell recovery after 7 days of acoustic stimulation. When low-frequency sounds of 100 Hz and 480 Hz were applied for 1 hour daily for four consecutive days, fibroblast migration increased by about 25% and 55%, respectively.

This study was designed around the hypothesis that the mechanical aspects of sound—especially vibration amplitude—may influence how skin cells respond to external stimuli. Because lower-frequency acoustic signals are known to generate stronger vibrations in both liquid and solid environments, we speculated that such physical oscillations might act as cues to support tissue repair. This study investigated the impact of different audible frequencies on skin cell migration by correlating vibration amplitude, measured using a 650 nm laser, with healing outcomes.

Methods

Human skin cell culture:

The Detroit 551 cell line, derived from human skin cells, was purchased from the Korea Cell Line Bank. Cells were grown in RPMI 1640 medium supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% penicillin–streptomycin (100 U/mL penicillin, 100 µg/mL streptomycin) to prevent microbial contamination. Cells were incubated at standard culture conditions (37°C, 5% CO_2). Cells were maintained in a healthy condition by changing to fresh media every three to four days.

Scratched wound healing assay:

After the skin cells were attached to the surface of the culture plate, a pipette tip was used to make a cross-shaped wound. Following scratch creation, the spent medium was aspirated to remove detached cells, and fresh culture medium was added to ensure consistent recovery conditions. Then, the fresh media was added. Bright-field images of the scratch area were captured immediately after wound creation (Day 0) to document the initial wound width. For each experimental condition, additional bright-field images were acquired after 4 days or 7 days of recovery to quantify cell migration and wound closure over time. All images were taken using identical microscope settings—including magnification, exposure, and field of view—to ensure consistent comparison across conditions.

Soundwave treatment:

Various frequencies were applied to the cultured cells using the soundwave generator: 100 Hz, 480 Hz, 1,000 Hz, 10,000 Hz, and 20,000 Hz. Cells were exposed to sound waves for one hour per day over four consecutive days. Sound waves were applied from above the culture plate to generate mechanical microvibrations.

Analyzing sound wave vibration and amplitude:

To measure the amplitude of the vibration caused by each frequency of sound, a 650 nm red laser module was utilized, along with a reflective surface placed at the bottom of the culture plate. The laser beam was incident at an oblique angle (approximately 45 degrees) to the surface of the culture plate. It was reflected by a small, lightweight reflective tape or mirror attached to the underside of the plate. The reflected beam was projected onto a white screen at approximately 1 meter, allowing for visualization of the beam displacement caused by micro-vibrations of the culture plate surface. When sound waves of different frequencies were applied above the culture plate, the reflected laser point vibrated measurably due to the induced vibrations. A high-resolution video camera was used to record the displacements at a rate of 60 frames per second. The peak-to-peak displacement of the reflected laser point was measured frame by frame using motion analysis software (e.g., ImageJ with the Manual Tracking plugin). The amplitude of vibration (in millimeters) was calculated as the total amplitude using the angular geometry of the arrangement and the distance between the reflective surface and the screen, obtained through trigonometric conversion of the observed displacement.

Results

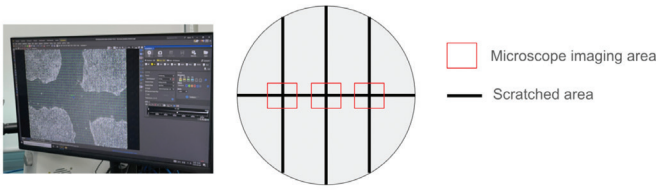


Figure 1: Schematic of Scratch Assay and Imaging Setup. The experimental setup used to evaluate wound closure in human dermal fibroblasts is shown. This figure illustrates how wound healing was assessed using a scratch assay. A uniform "wound" was created by dragging a pipette tip across a confluent monolayer of Detroit 551 fibroblasts, producing a gap where cells had been removed. Three fixed imaging sites along the scratch (red rectangles) were selected to ensure consistent measurements across time. Bright-field images were taken at Day 0 (immediately after scratching) and at later time points to track how quickly cells migrated into the wound area. This standardized setup allowed reliable comparisons of wound closure between control and sound-stimulated groups.

Figure 1 illustrates the experimental setup used to evaluate wound closure in cultured human dermal fibroblasts following scratch induction and soundwave treatment. The left panel depicts a live-cell imaging system in which confluent monolayers of Detroit 551 fibroblasts were scratched with a pipette tip to simulate wound gaps. The right panel shows a schematic of the culture dish, including the scratch regions (black vertical lines) and the three standardized imaging sites (red rectangles) selected along each scratch line. These positions were monitored using time-lapse microscopy to enable consistent and quantitative assessment of cell migration and wound healing

in response to microvibrations induced by audible sound stimulation.

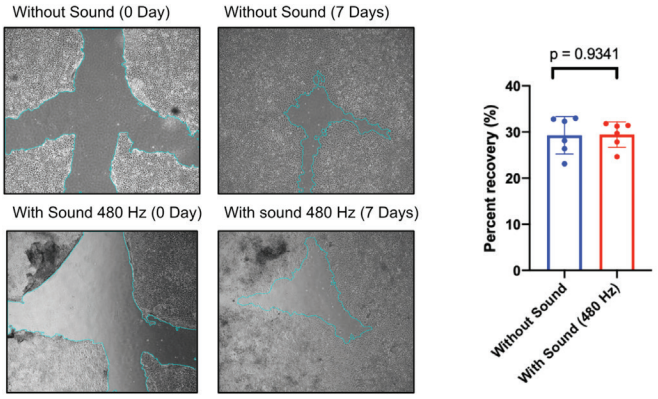


Figure 2: Effect of 480 Hz audible sound stimulation on wound healing over 7 days. Representative images of fibroblast scratch assays in control and 480 Hz sound-treated groups are shown at Day 0 and Day 7. After 7 days of daily 1-hour stimulation, the 480 Hz group showed a mean wound closure of 30.8% ± 4.1%, which was statistically indistinguishable from the control group (29.6% ± 3.7%). A two-tailed unpaired t-test confirmed that the difference between the two groups was not significant (p = 0.9341), indicating that prolonged 480 Hz exposure did not improve fibroblast migration or wound healing under the tested conditions (n = 6).

The impact of 480 Hz audible sound treatment on *in vitro* wound healing for 7 days is presented in Figure 2. As shown in the first row, control samples that were not sound-stimulated display clear boundaries of the scratch at Day 0, with partial wound closure apparent by Day 7. In comparison, the lower row displays the experimental group, which was exposed to 480 Hz sound for 1 hour per day. Clear wound boundaries are also visible on Day 0, and similar wound closure is observed at Day 7. To measure healing, the edges of the cells were highlighted in cyan, and the wound area was measured over time. The right bar graph indicates the percentage recovery of both conditions after 7 days. The average wound closure rate in the control (without sound) and the experimental (480 Hz) groups was about 30 percent. There was no significant difference between the two groups. These results indicate that audible sound stimulation at 480 Hz had no significant effect on wound healing in this model under the specified conditions. Thus, the two groups exhibited comparable levels of cell migration and wound closure over the 7 days.

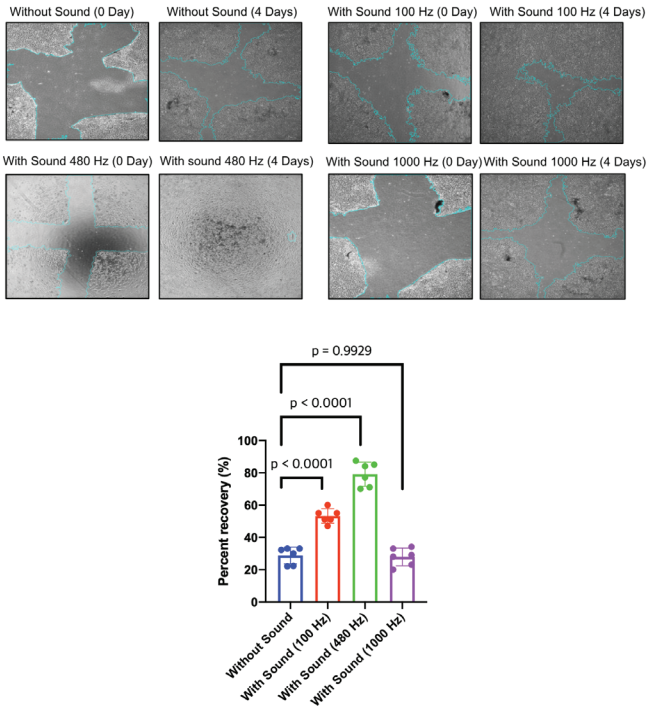


Figure 3: Frequency-specific effects of audible sound (100 Hz, 480 Hz, and 1,000 Hz) on wound healing over 4 days. Fibroblast scratch assays were analyzed after exposure to different frequencies for 1 hour daily over four days. These results demonstrate that wound-healing responses are highly frequency-dependent, with low-frequency audible sound—particularly 100 Hz and 480 Hz—producing the most robust enhancement of fibroblast migration (n = 6).

Figure 3 shows how different sound frequencies—specifically 100 Hz, 480 Hz, and 1,000 Hz—affected wound healing in human fibroblast cultures over four days. The images compare the initial scratch (day 0) and the state of healing after four days. Cells exposed to 100 Hz and 480 Hz demonstrated markedly greater wound closure compared to the untreated control, indicating a frequency-dependent enhancement of fibroblast migration. Cyan lines were drawn on each image to mark the wound edges and track changes over time. On the fourth day, the cultures treated with 100 Hz and 480 Hz sound showed noticeably better healing than the control group ($p<0.0001$), especially at 480 Hz. The 1,000 Hz group, on the other hand, looked very similar to the control ($p = 0.9929$). This indicates that the healing response may depend on the frequency of the sound, with 480 Hz being the most effective among the three tested.

DOI: 10.36838/IJHSR88.44

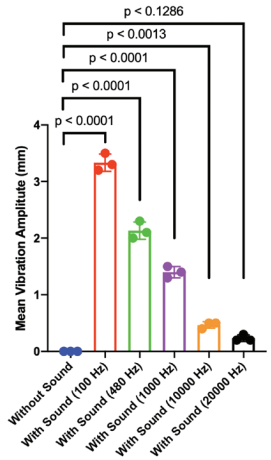


Figure 4: Frequency-dependent vibration amplitude induced by audible sound. Vibration amplitudes produced by sound frequencies ranging from 100 Hz to 20,000 Hz were measured using a 650 nm laser displacement tracking system. The strongest vibrations occurred at 100 Hz (3.4 mm) and 480 Hz (2.8 mm), both significantly higher than the control ($p < 0.0001$) ($n = 6$). A small reflective surface was attached under the culture plate, and a laser pointer was aimed at it at a shallow angle. As sound waves caused the plate to vibrate, the reflected laser dot moved on a screen positioned ~1 meter away. High-speed video recordings captured this movement, and peak-to-peak displacement was quantified using image analysis software. The bar graph shows that low-frequency sounds (100 Hz and 480 Hz) produced the largest vibration amplitudes (3.4 mm and 2.8 mm), while higher frequencies generated progressively smaller displacements. These mechanical differences help explain why only the low-frequency groups showed enhanced wound healing.

Figure 4 illustrates the vibration amplitudes generated by different audible sound frequencies (100 Hz to 20,000 Hz), as measured using a laser-based tracking setup with a wavelength of 650 nm. In the absence of sound, vibration was virtually undetectable, confirming that no external mechanical stimulation was applied to the control group—offering a reliable baseline. The most pronounced vibration occurred at 100 Hz (3.4 mm), followed by 480 Hz (2.8 mm), both of which were significantly stronger than the control and reflect the system's high sensitivity to low-frequency auditory input. Above 480 Hz, the amplitude gradually declined: approximately 1.6 mm at 1,000 Hz, 0.5 mm at 10,000 Hz, and just 0.2 mm at 20,000 Hz. Statistical analysis confirmed that vibration amplitudes at 100 Hz and 480 Hz were significantly higher than the control ($p < 0.0001$), and all other frequencies—except 20,000 Hz—also demonstrated statistically significant increases compared to the control. The amplitude at 20,000 Hz, however, was not statistically different from the control. Above 480 Hz, vibration intensity consistently declined as frequency increased. Interestingly, the most noticeable improvements in cell migration occurred at 100 Hz and 480 Hz—both frequencies that produced the strongest microvibrations—pointing to a possible link between vibration strength and the wound-healing effects of sound. Taken together, these findings indicate that low-frequency audible sound induces stronger mechanical stimulation, which in turn plays a key role in promoting fibroblast migration and wound repair.

Discussion

Fibroblasts exposed to 480 Hz audible sound for one hour daily over 4 days exhibited approximately 55% greater wound closure, indicating a frequency-specific enhancement of healing. Previous results showed that continuous sound exposure for seven days does not enhance healing and, in some cases, such as at 480 Hz, suppresses recovery. Thus, precise control of amplitude and exposure is needed.

To address this, we used brief daily hour sound exposures for four days and measured vibration amplitude with a 650 nm laser. The goal was to determine how sound-induced vibrational forces affect cellular behavior, particularly in the context of migration and repair.[7]

The selection of 100 Hz and 480 Hz as key test frequencies was based on both theoretical and empirical considerations from prior acoustic-biology research. Low-frequency audible sound (<500 Hz) is known to generate stronger mechanical displacements in solids and liquids due to reduced damping, making this range ideal for producing measurable microvibrations that cells can detect.[1–3] Previous studies also reported enhanced fibroblast migration around ~100 Hz and improved cytoskeletal remodeling under low-frequency mechanical cues, suggesting a biologically responsive window in this range.[3,14] In addition, preliminary pilot tests in our system showed that vibration amplitudes at 100–500 Hz were markedly higher than those at higher frequencies, providing a practical justification for selecting these points for deeper investigation. The inclusion of 480 Hz specifically allowed us to test a mid-low frequency that still produced substantial mechanical stimulation but had not been widely studied in wound-healing literature, enabling a comparative analysis across different parts of the audible spectrum. Frequencies above 1,000 Hz and into the ultrasonic range were included primarily as mechanistic controls to examine how diminishing vibration amplitude corresponds to reduced biological effects. Overall, this strategy enabled a frequency-resolved evaluation across the audible range while focusing on low-frequency candidates most likely to generate strong mechanobiological responses.

Building on these observations, we next evaluated how 100 Hz and 480 Hz sound stimulation influenced fibroblast migration and wound closure within our scratch assay model. The group exposed to 480 Hz showed about 55% more healing than the control group ($p < 0.0001$), suggesting that this frequency might help cells move more effectively. In contrast, higher frequencies, such as 1,000 Hz, showed no statistically significant benefit, corroborating previous observations that regenerative responses are more prominent at lower frequency ranges, potentially due to the sensitivity of the actin cytoskeleton.[1,2]

These findings indicate that cellular responses are closely linked to the magnitude of mechanically induced microvibrations, suggesting that vibration amplitude is a key determinant of the observed migratory enhancement. By better understanding this relationship, we may be able to enhance the use of sound in treatments that support tissue repair and regeneration.

While most past research has used ultrasound or very low-frequency sounds, our method employed regular sound in the air, eliminating the need for special tools or gels.[7-16] These

characteristics suggest that the method could be broadly applicable as a low-cost, non-invasive therapeutic approach—especially in underserved communities and among vulnerable populations such as older adults or diabetic patients.[13]

Although the findings are encouraging, one important limitation is that molecular-level changes were not directly assessed in this study. However, earlier research indicates that acoustic vibrations can stimulate core mechanotransduction cascades, in which physical forces from sound are translated into biochemical signals within the cell. In particular, the YAP/TAZ pathway is a well-established mechanism that detects microvibrations through cell structures and responds by moving these regulators into the nucleus, ultimately promoting the expression of genes related to migration, proliferation, and extracellular matrix remodeling—key processes in wound healing.[12] Fibroblast responses to mechanical signals often involve ERK and PI3K/Akt pathways, which also mediate cytoskeletal changes and cell movement.[17] This suggests exposure to 480 Hz sound may activate such mechanotransduction pathways through specific vibrational cues. Our study employed a simple scratch test and examined only one cell type; additional cell types and direct analysis of cell signaling remain unaddressed.

A key limitation of this study is the exclusive use of a single human dermal fibroblast line (Detroit 551), which restricts the generalizability of the findings to other skin-relevant cell types, such as keratinocytes, endothelial cells, or immune cells that also contribute to wound repair. Because fibroblasts represent only one component of the wound-healing environment, the observed benefits of 100 Hz and 480 Hz microvibrations may not fully reflect how entire tissues respond to audible sound stimulation. Additionally, all experiments were conducted *in vitro*, where the mechanical properties of culture plastics differ substantially from those of living tissue, and thus, *in vivo* confirmation is essential to determine whether similar vibration amplitudes can be safely and effectively delivered in real wounds. Longer-term safety data are also lacking: although brief, daily hour exposures were well tolerated *in vitro*, prolonged or repeated vibrational stimulation in living organisms may pose risks such as inflammation, tissue fatigue, or unwanted mechanotransduction effects. Consequently, future studies should incorporate multiple human cell types, 3D or organoid skin models, animal wound-healing studies, and systematic biocompatibility assessments to evaluate both efficacy and safety before translation to clinical applications.

Although we did not examine the cells in this study, earlier research indicates that sound-induced vibrations may trigger key pathways, such as YAP/TAZ, ERK, or PI3K/Akt. These are involved in cell movement and growth, which could explain the improved healing seen with certain frequencies. Future experiments—like western blotting, gene expression studies, or immunofluorescence—could help identify which signals are active. Animal models would also help test how well this works in real tissues and whether it's safe for longer-term use.

In this context, the measured vibration amplitudes help explain how audible sound may couple into specific mechanotransduction pathways. The 100 Hz and 480 Hz conditions generated the largest plate displacements (3.4 mm and 2.8 mm, respectively), indicating that fibroblasts in these groups experienced greater substrate strain than at higher frequencies. Such cyclic micro-deformations are known to be sensed primarily through integrins and focal adhesion complexes, which transmit forces to the actin cytoskeleton and upstream regulators of YAP/TAZ, ERK, and PI3K/Akt signaling.[6,8,10,15–17] When the deformation is large enough—yet still within a non-damaging range—these pathways are activated, driving nuclear translocation of YAP/TAZ and phosphorylation of ERK/Akt, which in turn upregulate genes involved in cytoskeletal remodeling, migration, and extracellular matrix production. Conversely, the smaller amplitudes observed at ≥1,000 Hz likely generate sub-threshold mechanical cues that are insufficient to robustly engage these force-sensitive signaling networks, consistent with the minimal change in wound closure at 1,000 Hz and 20,000 Hz. Taken together, our data support a model in which vibration amplitude, rather than frequency per se, is the critical physical parameter that determines whether audible sound activates mechanotransduction cascades that promote fibroblast migration and wound repair.

Conclusion

The study shows that low-frequency sound (100 Hz and 480 Hz) causes critical microvibrations in the experimental system. This also suggests a high possibility of promoting mechanotransduction in cellular conditions that facilitate wound healing. Conversely, increasing frequency above 1,000 Hz yielded successively smaller vibration amplitudes, and little mechanical response was recorded at ultrasonic frequencies (10,000 to 20,000 Hz). Our study results confirm that low-frequency sound induces a more powerful mechanical stimulus, significantly improving fibroblast migration and wound closure. This suggests the potential of audible sound as a noninvasive and accessible tool in regenerative medicine. The inverse correlation between frequency and vibration amplitude observed above 480 Hz confirms the hypothesis that audible sound can be used as a non-invasive mechanical stimulus to induce positive cellular effects, including migration and proliferation. These results suggest that low-frequency audible sound may serve as a practical, noninvasive mechanical stimulus capable of enhancing fibroblast migration, highlighting its potential utility in regenerative medicine applications. Going forward, studies should investigate how vibration-induced effects relate to cellular changes—such as cytoskeletal dynamics, molecular signaling, and gene activity—to better understand which sound frequencies are most effective and why.

Acknowledgments

I would like to thank my mentor, Dr. Woo Rin Lee, for being a steady source of support throughout this project. I built this experiment based on my earlier keratin imaging work. When I had trouble measuring vibrations, Dr. Lee suggested simple adjustments—such as adjusting the angle or the different surface—that helped me obtain clear results.

DOI: 10.36838/IJHSR88.44

References

1. Armand, A. C.; Bikaran, M.; Gardner, T. B.; Matthew, M. K. The role of infrasound and audible acoustic sound in modulating wound healing: A systematic review. Int. Wound J. **2025**, 22 (5), e70243. https://doi.org/10.1111/iwj.70243.
2. Brugger, M. S.; Baumgartner, K.; Mauritz, S. C. F.; Gerlach, S. C.; Röder, F.; Schlosser, C.; Fluhrer, R.; Wixforth, A.; Westerhausen, C. Vibration-enhanced cell growth induced by surface acoustic waves as an *in vitro* wound-healing model. Proc. Natl. Acad. Sci. U.S.A. **2020**, 117 (50), 31603-31613. https://doi.org/10.1073/pnas.2005203117.
3. Brugger, M. S.; Schnitzler, L. G.; Nieberle, T.; Wixforth, A.; Westerhausen, C. Shear-horizontal surface acoustic wave sensor for non-invasive monitoring of dynamic cell spreading and attachment in wound healing assays. Biosens. Bioelectron. **2021**, 173, 112807. https://doi.org/10.1016/j.bios.2020.112807.
4. Stamp, M. E. M.; Brugger, M. S.; Wixforth, A.; Westerhausen, C. Acoustotaxis – *in vitro* stimulation in a wound healing assay employing surface acoustic waves. Biomater. Sci. **2016**, 4, 1092-1099. https://doi.org/10.1039/C6BM00125D.
5. del Rosario-Gilabert, D.; Valenzuela-Miralles, A.; Esquiva, G. Advances in mechanotransduction and sonobiology: Effects of audible acoustic waves and low-vibration stimulation on mammalian cells. Biophys. Rev. **2024**, 16, 783–812. https://doi.org/10.1007/s12551-024-01242-1.
6. Dupont, S.; Morsut, L.; Aragona, M.; Enzo, E.; Giulitti, S.; Cordenonsi, M.; *et al*. Role of YAP/TAZ in mechanotransduction. Nature 2011, 474 (7350), 179–183. https://doi.org/10.1038/nature10137. 7. Ngo, O.; Niemann, E.; Gunasekaran, V.; Sankar, P.; *et al*. Development of Low Frequency (20-100 kHz) Clinically Viable Ultrasound Applicator for Chronic Wound Treatment. IEEE Trans Ultrason Ferroelectr Freq Control. **2019**, 66 (3), 572-580. https://doi.org/10.1109/TUFFC.2018.2836311.
8. Ingber, D. E. Cellular mechanotransduction: Putting all the pieces together again. FASEB J. **2006**, 20 (7), 811–827. https://doi.org/10.1096/fj.05-5424rev.
9. Kumeta, M.; Takahashi, D.; Takeyasu, K.; Yoshimura, S. H. Cell type–specific suppression of mechanosensitive genes by audible sound stimulation. PLoS. One. **2018**, 13 (1), e0188764. https://doi.org/10.1371/journal.pone.0197307.
10. Liu, Y.; Pan, X. Y.; Zhang, X. X.; Sun, J. L.; Mao, Y. H.; Yang, Y.; Wei, Z. T. Role of mechanotransduction mediated by YAP/TAZ in the treatment of neurogenic erectile dysfunction with low-intensity pulsed ultrasound. Andrology. **2023**, 11 (7), 1514–1527. https://doi.org/10.1111/andr.13438.
11. Mohammed, T.; Murphy, M. F.; Lilley, F.; Burton, D. R.; Bezombes, F. The effects of acoustic vibration on fibroblast cell migration. J. Biomech. **2016**, 49 (14), 3360–3366. https://doi.org/10.1016/j.msec.2016.07.037.
12. Panciera, T.; Azzolin, L.; Cordenonsi, M.; Piccolo, S. Mechanobiology of YAP and TAZ in physiology and disease. Nat Rev Mol Cell Biol. **2017**, 18 (12), 758-770. https://doi.org/ 10.1038/nrm.2017.87.
13. Wang, C.; Shirzaei Sani, E.; Gao, W. Wearable bioelectronics for chronic wound management. Adv. Funct. Mater. **2021**, 32 (17), 2111022. https://doi.org/10.1002/adfm.202111022.
14. He, Y.; Xia, J.; Mai, J. D. H.; Upreti, N.; *et al*. Acoustic technologies for the orchestration of cellular functions for therapeutic applications. Sci. Adv. **2025**, 11 (29), eadu4759. https://doi.org/10.1126/sciadv.adu4759.
15. Rosario-Gilabert, D. D.; Valenzuela-Miralles, A.; Esquiva, G. Advances in mechanotransduction and sonobiology: effects of audible acoustic waves and low-vibration stimulations on mammalian cells. Biophys. Rev. **2024**, 16 (6), 7830812. https://doi.org/10.1007/s12551-024-01242-1.
16. Crozet, F.; Levayer, R. Emerging roles and mechanisms of ERK pathway mechanosensing. Cell. Mol. Life. Sci. **2023**, 80 (12), 355. https://doi.org/10.1007/s00018-023-05007-z.
17. Jackson, B. T.; Artlett, C. M. Fibroblast migration in fibrosis. Fibrosis. **2025**, 3(2), 10009. https://doi.org/10.70322/fibrosis.2025.10009.

Author

Cheng Min Sun is a student researcher with a strong interest in physics and cellular biomechanics. This project reflects Cheng Min's broader commitment to exploring innovative, non-invasive methods for enhancing cellular responses and advancing regenerative medicine.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Exploring the Factors that Influence Gen Z's Trust Toward and Intentions to Use AI in Daily Life

Sebahattin Polat

TEV Inanc High School, 25, 4126th Street, Muallimkoy Neighborhood, P.O. Box 125, Gebze, Kocaeli, 41400, Turkey; sebo.polatt@gmail.com
Mentor: Ceyda Maden-Eyiusta

ABSTRACT: This study examines the psychological and social factors shaping Generation Z's trust in and intention to use Artificial Intelligence (AI) technologies in daily life, applying the Unified Theory of Acceptance and Use of Technology (UTAUT) framework. Survey data from 87 high school students in Istanbul, Turkey, reveal that performance expectancy (perceived usefulness) and attitude toward AI emerge as the strongest predictors of both trust in AI and behavioural intention to adopt it. Mediation analyses confirm that attitude functions as a key psychological mechanism linking performance expectancy to trust and usage intention. In contrast, effort expectancy (perceived ease of use) and facilitating conditions show limited or no direct effects, suggesting that for digitally native youth, perceived value outweighs usability concerns. Peer influence is found to shape attitudes but does not directly predict trust or usage intention. These findings underscore the importance of focusing on value-driven design and trust-building strategies when developing AI systems for younger users. The study contributes to both UTAUT literature and AI adoption research by providing a generational lens on technology acceptance, highlighting how digitally native Generation Z users prioritize perceived usefulness and trust over usability and infrastructural support when engaging with AI technologies.
KEYWORDS: Systems Software, Human/Machine Interface, Gen Z, Trust, Artificial Intelligence.

Introduction

Artificial Intelligence (AI) is a tool that makes human life easier by enabling machines to perform tasks that require human intelligence and skills, thereby transforming various aspects of human life and conducting research. It is also described as the simulation of human intelligence processes by machines, especially computer systems. These processes primarily involve learning, reasoning, and self-correction.[1] AI may replicate human skills in many different areas, such as decision making, pattern recognition, problem solving, and even natural language understanding.[2] Unlike ordinary software, which can be programmed to exhibit intelligent behavior, AI-guided systems can learn from data, adapt to new information, and operate autonomously without explicit human intervention. Accordingly, AI agents, which refer to advanced generative AI systems designed to perform goal-driven tasks autonomously, are becoming an integral part of our lives.

A fundamental difference between AI and other technologies lies in its ability to function as a 'black box' system, where the internal decision-making processes are often opaque and difficult to interpret.[3] This opacity becomes especially consequential as AI systems engage users through natural-language interactions that simulate human communication. Although these capabilities offer more seamless and user-friendly experiences, they simultaneously intensify concerns about transparency and accountability.[4] Many AI models, especially those based on deep learning and information processing, can provide outputs that even their developers cannot fully interpret. In other words, due to their over-parameterized black-box nature, it is often hard for individuals to understand the predictive outcomes of deep models.[5]

Despite these challenges, AI—particularly Generative AI (GenAI), which refers to AI systems capable of creating text, images, code, or other content in response to user input—has gained notable popularity and widespread adoption in recent years. This growing interest is especially prevalent among younger generations. Generation Z, in particular, is known for their motivation, willingness, and ability to adopt and utilize emerging technologies, including AI-based systems, as they are the first generation to be born and raised in a fully digital world.[6] Recent research further supports this observation: For example, 86% of Generation Z consider technology a crucial part of their daily lives, with many actively utilizing AI tools for both professional and personal purposes.[7] Additionally, 79% of Gen Z express a clear preference for AI-driven digital experiences, underscoring their deep integration with and fluency in emerging technologies.[8] Their dynamic interaction with tools like ChatGPT illustrates not only their comfort with such systems but also their potential to shape the future of human–AI collaboration.[9] Hence, it is reasonable to suggest that Generation Z demonstrates a higher level of technological engagement compared to preceding generations.

Understanding how individuals perceive and adopt AI requires insights from established models of technology adoption. The Unified Theory of Acceptance and Use of Technology (UTAUT), formulated by Venkatesh *et al.*[10], is one of the most widely used models for predicting the adoption and use of technology. The UTAUT aims to explain user intentions to use an information system and subsequent usage behavior

DOI: 10.36838/IJHSR88.50

by identifying four key determinants: performance expectancy (i.e., the degree to which an individual believes that using the system will help them attain benefits in job performance), effort expectancy (i.e., the perceived ease of use associated with the system), social influence (i.e., the extent to which individuals perceive that some important others believe they should use the system), and facilitating conditions (i.e., the degree to which an individual believes that an adequate technical infrastructure exists to support the use of the system).

Given their digital fluency and early exposure to emerging technologies, Generation Z users may evaluate these determinants differently from previous generations. They might develop a certain level of trust in AI systems more rapidly and intuitively. As such, applying UTAUT in the context of Gen Z's interaction with AI—especially GenAI tools—can shed light on this generation's adoption patterns and expectations, which might indeed shape the future direction of AI integration in both organizational and everyday contexts. Furthermore, exploring how primary UTAUT dimensions—such as performance expectancy, effort expectancy, social influence, and facilitating conditions—relate to Gen Z's trust in, attitudes toward, and intentions to use AI can offer valuable insights into how cognitive and social factors shape the acceptance of intelligent systems among digital-native users.

1.2. Research Focus:

Despite the growing presence of AI in everyday life and its rapid expansion across a wide range of domains,[11] there is still a limited understanding of the factors—and their relative influence—that shape Generation Z's adoption of and trust in AI technologies. This gap is particularly observable in generative AI tools, which are becoming deeply embedded in both educational and professional settings.[12,13]

By utilizing the UTAUT framework, this study aims to provide empirical insights into the psychological, social, and technological determinants that shape trust in AI and the adoption of AI systems among young users, with a particular focus on members of Gen Z. We expect that performance expectancy (i.e., the belief that using AI will enhance performance outcomes) and effort expectancy (i.e., the perceived ease of use) may be especially relevant for Generation Z, who tend to prefer smart, accessible, and efficient tools that integrate smoothly into their daily lives.[14] Moreover, social influence is likely to be particularly strong among this generation, given the influential role of peer networks, online communities, and social media in shaping their perceptions of new technologies.[15,16] Finally, facilitating conditions—such as access to necessary infrastructure, training, and technical support—may further shape Gen Z's willingness to adopt and trust AI tools in various life domains. In addition to the core UTAUT constructs, this study incorporates trust and attitude toward AI to better capture Gen Z's evaluative and relational engagement with AI systems. The autonomous and opaque nature of AI—particularly generative models—requires users to rely on outputs they cannot fully verify, making trust a central determinant of adoption. Attitude, as an individual's overall evaluative stance toward using AI, further complements the original UTAUT model by accounting for affective and experiential dimensions that are especially salient for digital-native users. Including these constructs, hence, extends the explanatory power of UTAUT and allows for a more comprehensive assessment of the mechanisms that shape Gen Z's intention to use AI. Figure 1 demonstrates the conceptual model of our study.

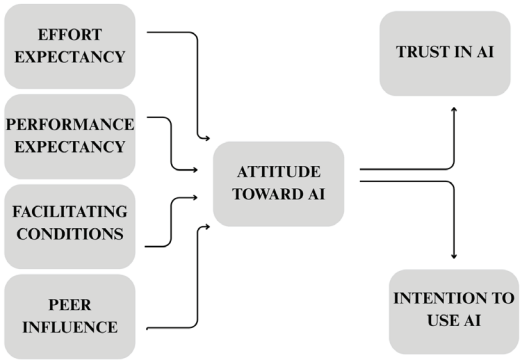


Figure 1: Conceptual model of the study. The model illustrates the relationships between effort expectancy, performance expectancy, facilitating conditions, and peer influence in shaping Generation Z's attitude toward AI (trust toward and intentions to use). It visually summarizes the hypothesized pathways tested in the study based on the UTAUT framework.

This study contributes to the growing conversations on technology adoption and user trust by exploring the factors that shape Generation Z's relationship with AI. By identifying the key drivers and barriers that influence Gen Z's trust in and adoption of AI, we offer timely insights into how this digitally native generation engages with emerging technologies. Importantly, by incorporating trust and attitude factors into the UTAUT framework, this study extends existing theoretical models to better account for the cognitive, affective, and relational dimensions of AI adoption among young users. As AI systems become increasingly integrated into education, work, and everyday decision-making, a clearer understanding of how Gen Z evaluates and adopts these tools is essential. The findings can inform the design of user-centered AI systems and guide policy efforts aimed at promoting the responsible and inclusive adoption of technology among younger users. In doing so, the study not only enriches the empirical base of the UTAUT framework but also responds to broader concerns about how trust in AI is built and sustained among young generations.

■ Methods

2.1. Sample and Data Collection:

Data were collected from 87 high school students in Istanbul, Turkey, by distributing a questionnaire through researchers' networks via a survey created from the Jotform application. Table 1 summarizes the key characteristics of the study sample. The table presents demographic and background information, including participants' age, gender, educational level, prior experience with AI technologies, interaction with artificial intelligence (AI) during a typical week, and the purposes for which they use AI.

Table 1: Sample Characteristics of the Participants. This table summarizes the demographic characteristics, grade levels, AI usage frequency, and primary purposes of AI use among participants.

Category	Details
Gender	• Female: 48 Participants (55.17%) • Male: 36 Participants (41.38%) • Prefer not to disclose: 3 Participants (3.45%)
Age	• Range: 14–20 • Mean age: 15.77 (SD = 1.43)
Grade Level	• 9th grade: 26 Participants (29.89%) • 10th grade: 20 Participants (22.99%) • 12th grade: 18 Participants (20.69%) • Preparatory class: 12 Participants (13.79%) • 11th grade: 11 Participants (12.64%)
Frequency of Weekly AI Interaction	• 5–6 times per week: 35 participants (40.23%) • Always interacting (6+): 23 Participants (26.44%) • Sometimes (3–4 times): 19 Participants (21.84%) • Rarely (1–2 times): 9 Participants (10.34%) • Never: 1 Participant (1.15%)
Purpose of AI Use	• Education and Learning: 70 Participants (80.46%) • Research and Assignment Completion: 68 Participants (78.16%) • Creative Writing and Content Creation: 28 Participants (32.18%) • Social Media and Communication: 20 Participants (22.99%) • Entertainment and Gaming: 16 Participants (18.39%) • Other: 13 Participants (14.94%)

2.2. Measures of Constructs:

All constructs were measured using well-established and validated scales from prior research, ensuring the reliability and construct validity of the study's measures. *Attitude* was measured using four indicators from Cao *et al.*[17] to assess the extent to which participants liked the idea of using AI, perceived it as a good or bad idea, and found using AI enjoyable (α = 0.78). *Effort expectancy*, defined as the perceived ease of using AI, was measured using four indicators (Cao *et al.*[17]; Venkatesh *et al.*[10] (α = 0.89). *Facilitating conditions* were measured using four items from Cao *et al.*[17] and Venkatesh *et al.*[10], assessing access to resources, AI compatibility with other technologies, and support availability for AI-related difficulties (α = 0.72). *Performance expectancy* was measured using five items from Cao *et al.*[17], assessing the extent to which individuals believe AI enhances their job performance (α = 0.85). *Peer influence* was measured using five indicators from Venkatesh *et al.*[10], adopted by Cao *et al.*[17] (α = 0.92). These indicators assess the extent to which participants feel encouraged to use AI based on the opinions of peers or other influential figures in their social environment. Three items from Cao *et al.*[17] measured participants' *intention to use AI in the future*, assessing their likelihood of using AI at work and frequency of use (α = 0.85). The *frequency and purpose of AI use* were measured using indicators developed by the researchers, assessing how often and for what purposes participants use AI in their daily lives. Finally, *trust in AI* was measured using twelve items adopted from Gulati, Sousa, and Lamas (2019), assessing reliability, competence, and perceived helpfulness. (α = 0.84)

2.3. Analytical Strategy:

Hierarchical regression analyses were conducted to examine the effects of the proposed predictors on behavioral intention to use and trust in AI. Three models were developed to test the hypothesized relationships. **Model 1** assessed the influence of effort expectancy (EE), performance expectancy (PE), peer influence (PI), and facilitating conditions (FC) on attitude toward AI. **Model 2** examined the effects of these four predictors, along with attitudes, on trust in AI. **Model 3** evaluated the impact of effort expectancy, performance expectancy, peer influence, facilitating conditions, and attitudes on the intention to use AI in the future.

■ Results and Discussion

3.1. Descriptive Statistics of Constructs:

Participants' responses to the questionnaire items were averaged to compute composite scores for each predictor—effort expectancy (EE), performance expectancy (PE), peer influence (PI), and facilitating conditions (FC)—as well as for the outcome variables. Table 2 presents the means, standard deviations, Cronbach's alpha coefficients (as indicators of internal consistency), and intercorrelations among all constructs. All statistical analyses were performed using SPSS version 15.

Table 2: Descriptive statistics and correlations. This table presents the means, standard deviations, internal consistency coefficients (Cronbach's α), and intercorrelations among all constructs. Performance expectancy showed the strongest positive correlations with attitude, trust, and intention to use AI, highlighting its central role in AI adoption. The correlations in Table 2 provide important insights into the relationships among key variables in the model. As expected, performance expectancy (PE) was strongly and positively correlated with both attitude toward AI (r = .49, p < .001), trust (r = .63, p < .001), and behavioral intention to use AI (r = .63, p < .001), underscoring its important role in shaping user openness to AI technologies. Additionally, attitude showed significant positive associations with trust (r = .51, p < .001) and intention to use (r = .59, p < .001), supporting its possible mediating role in the acceptance process.

	EE	PE	PI	FC	Attitude	Trust	Int. to Use	Age	Gender
EE Mean: 4.00 SD: 0.75	*0.89*								
PE Mean: 3.70 SD: 0.87	0.22*	*0.85*							
PI Mean: 3.56 SD: 1.09	0.11	0.35***	*0.92*						
FC Mean: 3.92 SD: 0.711	0.54***	0.11	0.01	*0.72*					
Attitude Mean: 4.18 SD: 0.57	0.24*	0.49***	0.40***	0.23*	*0.78*				
Trust Mean: 3.25 SD: 0.59	0.20	0.63***	0.31**	0.19	0.51***	*0.84*			
Int. to Use Mean: 3.94 SD: 0.87	0.20	0.63***	0.26*	0.23*	0.59***	0.45***	*0.85*		
Age Mean: 15.83 SD: 1.40	0.05	0.15	0.06	0.26*	-0.02	0.15	0.09	-	
Gender Mean: 0.62 SD: 0.55	-0.20	-0.02	-0.10	-0.29**	-0.16	-0.05	-0.06	-0.13	-
Grade/Class Mean: 1.95 SD: 1.36	-0.03	0.17	-0.11	0.14	-0.06	0.15	0.17	0.75***	-0.02

Notes: Cronbach's alphas are on the diagonal. *p < .05, **p < .01, ***p< .001

Effort expectancy (EE) showed smaller but significant associations with attitude (r = .24, p < .05), suggesting that ease of use is linked to overall positivity toward AI. However, facilitating conditions (FC) themselves did not show significant correlations with any of the outcome variables—namely, attitude, trust, or behavioral intention—implying that simply having access to resources may not be sufficient to influence these outcomes directly. Peer influence (PI) exhibited significant correlations with attitude (r = .40, p < .001), trust (r = .31, p < .01), and intention to use (r = .26, p < .05), emphasizing the relevance of social dynamics in Gen Z's AI adoption. Among the demographic variables, age was positively associated with facilitating conditions (r = 0.26, p < 0.05). At the same time, gender showed a negative correlation with the same variable

DOI: 10.36838/IJHSR88.50

(r = –0.29, p < 0.01), potentially indicating differences in perceived access or support. As expected, grade level correlated strongly with age (r = .75, p < .001), but showed no significant associations with the main constructs.

3.2. Regression Results:

The results of the regression models are reported below.

Model 1 (EE, PE, PI, and FC → Attitude): As shown in Table 3, Model 1 was statistically significant, $F(4, 82) = 8.01$, $p < .001$, explaining 28.1% of the variance in attitude ($R^2 = .281$, Adjusted $R^2 = .246$). Among the predictors, Performance Expectancy (PE; b = 0.166, t = 2.51, p = .014) and Peer Influence (PI; b = 0.149, t = 2.77, p = .007) emerged as significant positive predictors of attitude. These findings suggest that participants who believe AI is useful (in terms of enhancing their productivity, decision speed, and the likelihood of making important decisions) and are influenced by peers who use or advocate for AI are more likely to develop favorable attitudes toward its adoption. In contrast, Effort Expectancy (EE; b = 0.076, t = 0.87, p = .387) and Facilitating Conditions (FC; b = 0.124, t = 1.37, p = .174) did not significantly predict attitude.

Table 3: Regression analysis for predicting Attitude (Model 1). This table shows that performance expectancy and peer influence significantly predict a positive attitude toward AI, while effort expectancy and facilitating conditions do not have a significant effect.

Predictor	B	t	p
Intercept	2.23	5.69	< .001
Effort Expectancy	0.076	0.87	.387
Performance Expectancy	0.166	2.51	.014
Peer Influence	0.149	2.77	.007
Facilitating Conditions	0.124	1.37	.174

Model 2 (EE, PE, PI, FC, and Attitude → Trust): Model 2 was statistically significant, F(5, 81) = 9.83, p < .001, accounting for approximately 38% of the variance in trust ($R^2 = .378$, Adjusted $R^2 = .339$). However, the predictors did not contribute equally to the explanation of trust (please see Table 4). Only Performance Expectancy and Attitude emerged as significant positive predictors of Trust. This suggests that users who find AI technologies useful and have developed a favorable attitude toward them are more likely to trust them. Interestingly, Effort Expectancy, Peer Influence, and Facilitating Conditions did not significantly contribute to the prediction of trust when the other variables were accounted for.

Table 4: Regression analysis for predicting trust (Model 2). This table shows that performance expectancy and attitude significantly predict trust in AI, suggesting that both cognitive evaluations of usefulness and favorable attitudes are key drivers of trust.

Predictor	**B**	**t**	**p**
Intercept	0.586	1.30	.198
Effort Expectancy	-0.010	-0.12	.909
Perceived Usefulness	0.240	3.60	.001
Peer Influence	0.015	0.28	.781
Facilitating Conditions	0.079	0.88	.380
Attitude	0.343	3.19	.002

Model 3 (EE, PE, PI, FC, and Attitude → Intention to Use AI): Model 3 assessed the extent to which cognitive, social, and attitudinal factors predict individuals' intention to use AI technologies. Specifically, Effort Expectancy, Performance Expectancy, Peer Influence, Facilitating Conditions, and Attitude were entered as predictors. The model was statistically significant, F(5, 81) = 15.20, p < .001, and explained 48.4% of the variance in intention to use AI ($R^2 = .484$, Adjusted $R^2 = .452$) (please see Table 5). Among the predictors, Performance Expectancy (b = 0.390, t = 4.40, p < .001) and Attitude (b = 0.683, t = 4.78, p < .001) were strong and significant positive predictors of intention. These results indicate that individuals are more likely to intend to use AI when they perceive it as useful and hold favorable attitudes toward it. In contrast, Effort Expectancy (b = –0.147, p = .198), Peer Influence (b = –0.062, p = .394), and Facilitating Conditions (b = 0.210, p = .081) did not reach significance. Notably, Facilitating Conditions approached marginal significance (p = .081), suggesting a possible supporting role when infrastructure or support is available.

Table 5: Regression analysis for predicting trust (Model 3). In this table, intention to use AI is most strongly predicted by performance expectancy and attitude, with facilitating conditions approaching significance, while effort expectancy and peer influence are not significant predictors.

Predictor	B	t	P
Intercept	–0.384	–0.64	.523
Effort Expectancy	–0.147	–1.30	.198
Performance Expectancy	0.390	4.40	< .001
Peer Influence	–0.062	–0.86	.394
Facilitating Conditions	0.210	1.77	.081
Attitude	0.683	4.78	< .001

To investigate the underlying psychological mechanisms linking cognitive appraisals to trust and behavioral intention toward AI, we conducted two mediation analyses using the PROCESS macro for SPSS. Specifically, we tested whether Attitude mediates the relationship between predictor variables and (1) trust and (2) intention to use AI, using 5,000 bootstrap resamples and 90% confidence intervals.

Regarding Model 2, the results showed a significant indirect effect of Performance Expectancy on Trust mediated by Attitude. The indirect effect was 0.098, and the 90% confidence interval was [0.041, 0.164], excluding zero. This indicates that individuals who expect AI to enhance their performance are more likely to develop positive attitudes, which in turn increase their trust in AI technologies. The mediation effect was statistically meaningful and confirms that attitude serves as a psychological bridge between performance-related expectations and trust.

In the second model, Attitude also significantly mediated the relationship between Performance Expectancy and Intention to Use AI. The indirect effect was 0.195, with a 90% confidence interval of [0.128, 0.283], again clearly excluding zero. This suggests that the positive impact of PE on individuals' willingness to adopt AI is transmitted through enhanced attitudes, reinforcing the critical role of affective-cognitive appraisals in shaping behavioral intention.

■ Conclusion

Using the UTAUT framework, this study explored the psychological and social determinants that shape Generation Z's trust in and intention to use AI in daily life. The findings reveal that Performance Expectancy (PE) and Attitude are the most

important determinants of both trust in AI and behavioral intention to use it. Mediation analyses also confirmed that Attitude acts as a bridge through which PE influences trust and intention. In contrast, Effort Expectancy (EE) and Facilitating Conditions (FC) showed limited or non-significant effects across models. Moreover, Peer Influence (PI) was found to influence attitudes significantly, but not trust or intention to use, highlighting a more precise role in the adoption process.

Performance Expectancy emerged as the most influential determinant (regression analysis showed that PE had a $\beta = 0.39$, $p < .001$) in shaping how respondents perceived and interacted with AI technologies. The belief that AI enhances decision-making, productivity, and overall efficiency significantly contributed to both trust and behavioral intention. Attitude, defined as respondents' affective and evaluative orientation towards the use of AI, played an important mediating role. Participants who expected AI to improve their performance were more likely to develop a positive attitude, which in turn increased their trust in AI tools and their willingness to use them. These findings align with previous research and underscore the pivotal role of perceived utility in driving AI adoption, particularly among digitally native groups, such as Generation Z.

Unlike some previous studies, Effort Expectancy did not significantly predict trust, attitude, or intention. This finding may be explained by the high digital fluency of Generation Z participants, who typically interact effortlessly with technology. For them, ease of use is probably considered a basic expectation rather than a determining factor. Similarly, Facilitating Conditions, such as access to resources or support systems, had a limited impact, possibly due to the sample's high basic access to digital tools and infrastructure. These results suggest that among tech-savvy youth, the functionality and outcomes of AI are more important than accessibility or usability.

The study found that Peer Influence significantly influenced attitudes towards AI, but did not extend to trust or intention to use. This suggests that peers may shape initial perceptions and openness, but the decision to trust and adopt AI tools is ultimately more personal and performance-oriented. This pattern may reflect cultural or generational tendencies towards individual decision-making, as Generation Z often values autonomy in digital choices despite being socially interconnected. In addition, the limited influence of social norms may indicate that AI adoption is moving towards becoming a normative and expected behavior, reducing the relative weight of peer pressure.

The findings offer actionable insights for AI developers, educators, and policymakers aiming to foster trust and adoption among Generation Z. Efforts should prioritize demonstrating AI's value and performance benefits rather than focusing exclusively on its ease of use. Transparent, outcome-oriented systems aligned with users' goals, alongside educational initiatives highlighting AI's real-life applications, can strengthen both trust and engagement.

This study contributes to the UTAUT literature by providing a generational lens on technology acceptance, specifically for digital natives, such as Generation Z. In this context, traditional predictors such as Effort Expectancy and Facilitating Conditions carry less influence, while Performance Expectancy and Attitude play a central role. The mediating role of Attitude between perceived usefulness and trust/intention underscores the importance of emotional evaluations in Gen Z's adoption process. Future UTAUT applications should account for age-specific patterns and digital familiarity.

Although it provides valuable insights, the current study has several limitations. First, the sample size ($N = 87$) is below the recommended limit for large-scale generalizations, and participants were limited to high school students in Istanbul, Turkey, which limits the regional and cultural generalizability of the results. Furthermore, the age distribution is skewed towards the early high school years, which may limit the maturity and depth of participants' responses. Future research should include more diverse age groups, education levels, and cultural backgrounds, and adopt longitudinal designs to track how trust and intention to use AI develop as Generation Z progresses into higher education and professional life. Finally, the study did not incorporate other measures of participant background, such as duration of AI use or comfort with digital technologies, which could provide additional insights into adoption patterns; future research could incorporate these measures to better understand individual differences in AI engagement.

In conclusion, this study provides timely insights into the psychological and social factors shaping Generation Z's trust in and adoption of AI technologies. The findings demonstrate that for Gen Z, perceived usefulness—not ease of use or infrastructure—is the primary driver of both trust and intention to use AI. This shift underlines changing priorities in technology adoption among younger generations. As AI becomes increasingly integrated into education, career development, and daily life, it is essential to align system design with Gen Z's values, focusing on clear value creation, trust-building, and outcome-oriented communication. Empowering this generation with purposeful, transparent, and value-driven technologies will not only encourage adoption but also support the ethical and sustainable integration of AI across society.

Acknowledgments

I would like to sincerely thank Associate Professor Ceyda Maden-Eyiusta from the Faculty of Business at Ozyegin University for her valuable guidance, encouragement, and support throughout this research. I, as the corresponding author, greatly appreciate all the help received during this study.

References

1. Lifewire. *What is artificial intelligence (AI)?* 2021. https://www.lifewire.com/what-is-artificial-intelligence-5119206
2. Dwivedi, Y. K.; Hughes, L.; Ismagilova, E.; Aarts, G.; Coombs, C.; Crick, T.; Williams, M. D. *Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy*. Int. J. Inf. Manage. 2021, 57, 101994. https://www.sciencedirect.com/science/article/abs/pii/S026840121930917X
3. Rai, A. *Explainable AI: From black box to glass box*. J. Acad. Mark. Sci. 2020, 48(1), 137–141. https://link.springer.com/article/10.1007/s11747-019-00710-5

DOI: 10.36838/IJHSR88.50

4. Bulletin of the Atomic Scientists. *Why nobody can see inside AI's black box*. 2025. https://thebulletin.org/2025/01/why-nobody-can-see-inside-ais-black-box/
5. Li, X.; Xiong, H.; Li, X.; Wu, X.; Zhang, X.; Liu, J.; Bian, J.; Dou, D. *Interpretable Deep Learning: Interpretation, Interpretability, Trustworthiness, and Beyond*. arXiv, 2021. https://arxiv.org/abs/2103.10689
6. McKinsey & Company. *What is Gen Z?* 2024. https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-gen-z
7. CTA. *Exploring Gen Z views and preferences in technology.* Consumer Technology Association, 2024. https://www.cta.tech/press-releases/cta-research-exploring-gen-z-views-and-preferences-in-technology
8. NextWaveGenZ. *50+ Statistics About Generation Z in 2024: Digital Habits*. 2024. https://nextwavegenz.com/facts-statistics/50-statistics-about-generation-z-digital-habits-edition/
9. The Times. *Gen Z students won't use ChatGPT - but not because it's cheating*. 2025. https://www.thetimes.com/uk/education/article/gen-z-students-wont-use-chatgpt-but-not-because-its-cheating-v8rffjlc0
10. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*. 2012;36(1):157–178. https://www.jstor.org/stable/41410412
11. OECD. *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. OECD Publishing, 2023. https://doi.org/10.1787/08785bba-en
12. Bick, A.; Blandin, A.; Deming, D. J. *The rapid adoption of generative AI*. National Bureau of Economic Research (Working Paper 32966), 2024. https://www.nber.org/system/files/working_papers/w32966/w32966.pdf
13. Miao, F.; Holmes, W.; Huang, R.; Zhang, H. *AI and education: A guidance for policymakers*. Unesco Publishing, 2021. https://unesdoc.unesco.org/ark:/48223/pf0000376709?locale=en
14. Francis, T.; Hoefel, F. *'True Gen': Generation Z and its implications for companies*. McKinsey & Company, 2018. https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/true-gen-generation-z-and-its-implications-for-companies
15. Hapsari, A. Y.; Sukandi, P.; Dalimunthe, G. P.; Lisdayanti, A.; Sumadhinata, Y. E.; Nilasari, I. The impact of social media influencers on consumer behavior: A comparative analysis of Generation Z and millennials life style. Int. J. Humanit. Educ. Soc. Sci. 2024, 3(5), 2559–2563. https://ijhess.com/index.php/ijhess/article/view/988
16. Turner, A. *Generation Z: Technology and Social Interest.* J. Individ. Psychol. 2015, 71, 103–113. https://doi.org/10.1353/jip.2015.0021
17. Cao G, Duan Y, Edwards JS, Dwivedi YK. Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making. *Technovation*. 2021;106:102312. doi:10.1016/j.technovation.2021.102312.
18. Gulati S, Sousa S, Lamas D. Design, development and evaluation of a human-computer trust scale. *Behaviour & Information Technology*. 2019;:1–17. doi:10.1080/0144929X.2019.1656779.

Author

Sebahattin Polat:

A highly motivated high school student from TEV Inanc High School in Turkey, currently completing a student exchange year in Portland, USA. Passionate about Business and Psychology, with strong aspirations to attend top-tier universities, pursue advanced studies, conduct innovative research, and make meaningful contributions to these fields through academic excellence, intellectual curiosity, and creative problem-solving.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Leveraging Clinical Data for Machine-Learning-Based Heart Detection

Arth Bhardwaj, Janani Sekar

1) St. Francis High School, 1885 Miramonte Ave, Mountain View, CA 94040, USA; arthbhardwaj1234@gmail.com
2) Harvard University, Cambridge, MA 02138, USA

ABSTRACT: This study uses machine learning to evaluate demographics, cholesterol, and cardiovascular outcomes. We investigated cholesterol's prognostic power in two settings: incident heart disease and death among cardiac patients. Two datasets were used: UCI Cleveland (n=303; yes/no heart disease) and Faisalabad Institute of Cardiology, Pakistan (n=299; longitudinal heart failure outcomes). We hypothesized that elevated cholesterol would better detect new heart disease cases than in predicting mortality among individuals already diagnosed. Logistic regression and multilayer perceptron models were trained; AUC, recall, accuracy, and precision were used for evaluation. In Cleveland, a cholesterol-only logistic regression predicted heart disease with 53% accuracy; multivariable models adding age, ST-segment depression, and maximum heart rate reached 85% accuracy (AUC 0.91). In Faisalabad, cholesterol predicted death with 78% accuracy; after adding demographics and comorbidities, the multilayer perceptron achieved 97% accuracy, and logistic regression 82%. Best logistic models indicated diabetics had 2.6× higher death risk and that mortality increased by ~62% per unit of cholesterol. This comparative study shows that machine learning improves cardiovascular risk assessment in data-limited settings and that risk variables vary across illness stages.

KEYWORDS: Medical and Health Sciences, Cardiology, Machine Learning, Cholesterol, Risk Stratification.

Introduction

Nearly 18 million people die annually from cardiovascular disease (CVD), a significant cause of death worldwide. Thus, identifying and understanding the primary CVD risk factors is essential.[1] Cholesterol is one of numerous biochemical indicators, although its role in CVD is well known. High LDL cholesterol causes arterial plaques and atherosclerosis. These factors reduce blood flow and increase the risk of myocardial infarction and stroke. However, HDL cholesterol removes excess cholesterol from the circulation, preventing cardiovascular events.[2–4]

The risk of cardiovascular disease is not limited to high cholesterol. Age, sex, blood pressure, BMI, smoking, diabetes, and physical exercise interact complexly to produce cardiovascular disease.[5–7] These effects are worsened by urban stress, inactivity, and poor diet.[8,9] Numerous studies have demonstrated that smoking and diabetes considerably increase cardiovascular morbidity and mortality.

Physicians now use AI and ML to analyze large datasets, improving diagnosis and patient outcomes.[10] Machine learning algorithms can handle multiple risk factors at once, detecting subtle, non-linear relationships that statistical models miss. In the UCI Cleveland Heart Disease dataset, one of the most widely used benchmarks, clinical and demographic characteristics predict heart disease well.[11,12] ML approaches can increase cardiovascular risk-assessment accuracy and personalization, expanding on conventional instruments like the Framingham Risk Score.[13] ML-driven models must balance interpretability, prediction performance, transparency, explainability, and clinical actionability.

This study on cholesterol's predictive role in cardiovascular events aims to improve risk classification by adding clinical and demographic characteristics.[14] We wanted to identify people at risk of heart disease, predict mortality in cardiovascular patients, and compare the predictive power of cholesterol levels in these two contexts to see if they were enough. We hypothesized that elevated cholesterol levels would detect incident heart disease better than fatal cases. Our novel strategy separates variables affecting illness, beginning from mortality progression, to better understand the intricate web of processes that lead to cardiovascular risk.

We tested our hypotheses with two datasets. First was the UCI Cleveland Heart Disease dataset, which included 303 American patients with yes/no heart disease results. The Pakistani Faisalabad Heart Failure dataset includes 299 patients' longitudinal data, including mortality outcomes. We examined demographic and endpoint prediction tendencies using these datasets. With logistic regression and MLP models, survival and classification results may be assessed. In addition to cholesterol levels, we examined demographics, vital signs, comorbidities, and lifestyle factors for predictive potential.

For illness incidence and death, multivariate models with age, sex, blood pressure, comorbidities, and other clinical factors outperformed cholesterol-only models. Contrary to our expectations, cholesterol alone was better at distinguishing patients with and without a history of mortality than at predicting incident heart disease. Once the disease has progressed, cholesterol and other risk factors may be strong prognostic indicators.

This study shows how machine learning may improve tailored predictions and how crucial multidimensional risk modelling is for cardiovascular treatment. The work improves our under-

DOI: 10.36838/IJHSR88.56

standing of cardiovascular disease risk factors. It emphasises the need to include clinical, demographic, and biochemical data in prediction models by distinguishing disease-causing variables from death-causing factors. Such methods can inform targeted therapies and tailored management regimens in diverse and resource-constrained healthcare settings. This study reinforces the therapeutic value of combining classical markers like cholesterol with cutting-edge computational methods to improve patient care.

Methods

We compared the Faisalabad Institute of Cardiology's Pakistani clinical dataset and the University of California, Irvine's Cleveland Heart Disease dataset for cardiovascular outcome prediction. Since both datasets included demographic and clinical information, machine-learning predictive modeling was possible. We tested logistic regression and multilayer perceptron (MLP) neural networks to predict heart disease and death. Comparing linear and nonlinear modeling methods allowed us to evaluate their ability to capture complex risk-factor–outcome relationships.

We evaluated all datasets for completeness before modeling. Preprocessing was relatively straightforward in the UCI Cleveland dataset because most variables were complete; however, a small number of cholesterol entries are recorded as 0, which may represent missing values in the original source, and is discussed as a limitation. To ensure predictive analysis validity in the more complex and heterogeneous Faisalabad dataset, similar filters were employed to maintain only complete records. For machine-learning training, one-hot encoding transformed categorical data, including sex, chest-pain type, smoking status, and comorbidities. Standardizing continuous variables—heart rate, blood pressure, cholesterol, and age—to zero mean and unit variance enhanced model convergence and interpretability.

We evaluated the models using an 80/20 train–test split to verify that the test set appropriately reflected fresh data. Class balance was preserved by stratifying this split by outcome to reduce performance-metric bias. We selected the most informative variables through iterative feature selection using domain expertise and performance-guided testing. To assess prediction accuracy, we examined cholesterol-only models and models that gradually added demographic and clinical variables—Table 1. Summary of models, input features, and predicted outcomes for each dataset.

Two main models were created. Logistic regression provided a linear baseline for coefficient interpretation and revealed predictor relevance. Multilayer perceptrons—feedforward neural networks with hidden layers—can express nonlinear interactions and complicated correlations.

We examined how well algorithms predicted outcomes across the two datasets and how much cholesterol alone affected mortality and heart-disease risk relative to a multidimensional feature collection. This approach showed how machine learning can improve cardiovascular risk assessment by evaluating the utility of additional demographic and clinical data.

Table 1: Comparison of input variables and performance metrics for logistic regression (LR) and multilayer perceptron (MLP) models across the Cleveland and Faisalabad cohorts. Models with more features significantly outperformed cholesterol-only baselines: in Cleveland, multivariable LR achieved 85% accuracy (AUC 0.91), and in Faisalabad, the MLP achieved 97 % accuracy for mortality prediction—showing the power of multidimensional clinical data.

Model	Dataset	Input Features	Predicted Outcome
Logistic Regression	UCI Cleveland	Cholesterol (mg/dL)	Heart Disease
Logistic Regression	UCI Cleveland	Age (years), Sex, Resting Blood Pressure (mmHg), Cholesterol (mg/dL), Maximum Heart Rate (bpm), ECG ST Segment Depression (mm), Resting ECG, ECG ST Slope Value, Chest Pain, Exercise Angina	Heart Disease
Logistic Regression	Faisalabad	Cholesterol (mg/dL)	Mortality
Logistic Regression	Faisalabad	Age (years), Sex, Urban, Marital Status, Resting Blood Pressure (mmHg), Cholesterol (mg/dL), Maximum Heart Rate (bpm), ECG ST Segment Depression (mm), Resting ECG, ECG ST Slope Value, Chest Pain, Exercise Angina	Mortality
Multilayer Perceptron (Hidden Layers: [20, 20])	Faisalabad	Age (years), Sex, Urban, Marital Status, Resting Blood Pressure (mmHg), Cholesterol (mg/dL), Maximum Heart Rate (bpm), ECG ST Segment Depression (mm), Resting ECG, ECG ST Slope Value, Chest Pain, Exercise Angina, Diabetes, Smoking, Allergies, Family History	Mortality

Table 1 summarizes the study's datasets—UCI Cleveland and Faisalabad—machine-learning models, input attributes, and anticipated outcomes. As shown, model complexity and feature richness affected predictive performance for incident heart disease in the UCI dataset and mortality in the Faisalabad dataset. Initial models with cholesterol as the only predictor tested cholesterol's independent effect on risk prediction.

The efficacy of cholesterol as a sole outcome discriminator was assessed in these preliminary analyses using logistic regression and MLP models.

Age, sex, blood pressure, maximum heart rate, electrocardiogram (ECG) readings, chest-pain type, exercise-induced angina, and comorbidities, including diabetes and smoking status, were then added to the models. Using this multivariable strategy to measure each component's predictive value helped us build more accurate risk models. To ensure interpretability, logistic regression coefficients were used to calculate odds ratios (ORs) for each predictor. Using a 2×2 contingency table, ORs were calculated to investigate the relationship between cholesterol and heart disease. In our calculation, A and B represent high-cholesterol groups with and without heart disease, and C and D represent normal-cholesterol groups with and without heart disease; the OR was computed as $(A \times D) / (B \times C)$. A larger value indicates a higher likelihood of heart disease associated with elevated cholesterol.

We tested neural-network models with one, two, or three hidden layers and varying neuron counts. A grid search revealed that a two-hidden-layer network with 20 neurons per layer and ReLU activation performed best. Classification networks were trained using the Adam optimizer and binary cross-entropy loss. Training was discontinued when validation performance stopped improving to avoid overfitting. Using

feature selection, model modification, and regularization, we maximized expected accuracy and model generalizability.

We learned more about the multidimensional characteristics' additive value and the intricate, likely nonlinear relationships between clinical variables and cardiovascular outcomes by comparing linear logistic regression and nonlinear MLPs.

Results and Discussion

We tested the predictive power of cholesterol and other clinical parameters using logistic regression (LR) and multilayer perceptron (MLP) models trained on the UCI Cleveland and Faisalabad datasets. The original baseline model recognized heart disease in the UCI Cleveland dataset using cholesterol as a predictor with 53% accuracy. Although cholesterol is a helpful biomarker, this suggests it is not a strong diagnostic indicator alone. To improve prediction performance, we added age, maximum heart rate, ST-segment depression, resting blood pressure, and additional clinical variables. With this larger set, LR accuracy reached 85%, demonstrating how multidimensional clinical data improve risk assessment. In addition, the extended LR model obtained an AUC = 0.91, indicating accurate and consistent discrimination across thresholds.

The MLP model, which captures nonlinear variable interactions, achieved 84% accuracy with AUC = 0.88. According to these findings, adding several relevant features to either a linear or nonlinear model enhances heart-disease prediction. LR provides interpretability via odds ratios and coefficients, while MLPs model complex interactions. Figure 1 illustrates the frequency distributions of key variables in the Cleveland dataset.

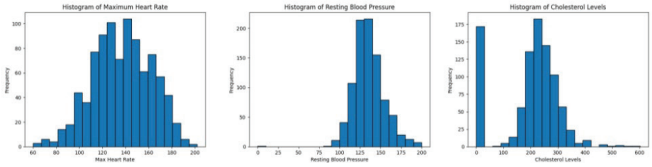


Figure 1: Histograms and density plots show distributions of key numeric variables in the Cleveland dataset. Because cholesterol(md/dl) alone yielded only 53% accuracy, these balanced distributions justify including additional variables such as age(years), ST-depression, and max heart rate, enabling an AUC of 0.91 with the multivariable LR model.

The cholesterol, resting blood pressure, and maximum heart rate histograms were approximately normal, while resting blood pressure showed a slight right skew. Skewed distributions can change the relative importance of variables in linear models; therefore, they should be examined carefully during model building and clinical interpretation. To ensure feature-representation uniformity and cross-dataset comparability, Figure 2 shows the Faisalabad dataset distributions. Visualizing these distributions helps explain population-specific demographic and clinical factors that affect model generalizability. Overall, these findings suggest that multidimensional machine-learning methods are beneficial in clinical practice and that adding clinical characteristics to cholesterol enhances cardiovascular risk prediction.

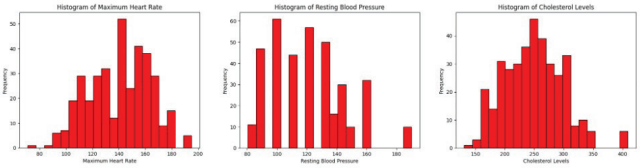


Figure 2: Visualization of key feature distributions in the Faisalabad cardiology dataset. Cohort differences visible here correspond to markedly higher mortality-prediction accuracy (up to 97%) once comorbidities and demographics are included, underscoring the impact of richer clinical features on prognostic models.

In the Faisalabad dataset, cholesterol-only models predicted mortality outcomes with 78% accuracy in cardiovascular disease patients. Adding demographic and comorbidity-related factors such as age, diabetes, smoking, and urban dwelling improved predictive performance. By combining multidimensional clinical data, LR achieved 82% accuracy. The MLP captured complex nonlinear interactions among variables with 97% accuracy, as seen in Figure 7. These findings show that machine learning can improve prognostic evaluations beyond single-variable models and that high-risk populations require comprehensive feature sets.

Outcome-segregated feature distributions support these findings. Figures 3 and 4 show that deceased patients were older and had higher cholesterol, while survivors tended to have higher maximum heart rates. These trends imply that cholesterol is a meaningful predictor on its own but becomes considerably stronger when combined with demographic and clinical characteristics. These results suggest that neural networks and other multidimensional modeling methods can improve risk classification and personalized treatment regimens by predicting poor cardiovascular outcomes in diverse patient groups. Importantly, these visual shifts are clinically plausible and help interpret why certain predictors receive larger learned weights in the multivariable models.

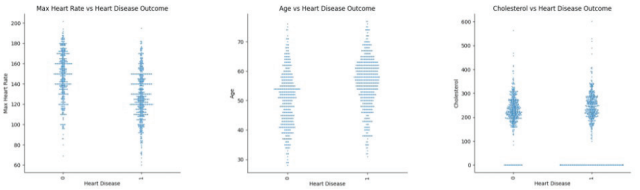


Figure 3: Boxplots showing how age (years), cholesterol (mg/dl), and max heart rate (bpm) differ between healthy and diseased patients. Individuals with disease are generally older, have higher cholesterol, and lower max heart rate—patterns that allowed LR to improve from 53 % (cholesterol-only) to 85 % accuracy with multiple predictors.

Figure 4: Heart-rate(bpm) response, matching the model's learned coefficients. Diabetes increased risk approximately 2.6×, explaining the superior performance of multifeature MLP models.

DOI: 10.36838/IJHSR88.56

To understand variable relationships, we examined correlations with age (Figures 5 and 6). We evaluated the statistical significance of individual predictors and model performance across outcome groups using one-way ANOVA for continuous variables (cholesterol, age) and chi-square tests for categorical variables (smoking, diabetes). Paired t-tests compared cholesterol-only and multivariable configurations across five cross-validation rounds. All tests were two-tailed with $\alpha = 0.05$. Including demographic and clinical parameters beyond cholesterol significantly improved prediction performance in both LR ($p = 0.002$) and MLP ($p < 0.001$).

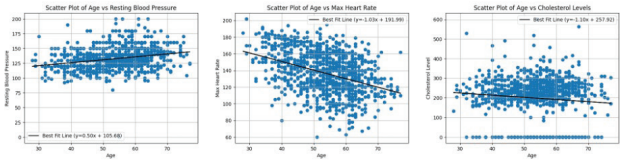


Figure 5: Scatter plots showing relationships of age with resting blood pressure (mmHg, cholesterol (mg/dl), and max heart rate (bpm). The upward trend in BP/cholesterol and decline in heart rate with age produced an additive predictive signal, helping LR reach AUC 0.91 when age was combined with ECG and exercise-related features.

Figure 6: Analogous age-related trends for the Faisalabad cohort. The synergy of age with comorbidities such as diabetes and smoking strengthened mortality prediction, explaining why nonlinear MLPs achieved the highest prognostic accuracy.

An odds-ratio (OR) analysis assessed these links. Diabetics had 2.61× higher risk of death, and mortality increased by ~62% per unit of cholesterol. In addition, higher ECG ST-segment slope values approximately doubled to tripled the risk of death. In this cohort, top mortality models had precision and recall above 0.90, supporting reliable identification of high-risk patients. In clinical practice, risk markers beyond cholesterol should therefore be considered when making prognostic predictions.

Comparison with prior work: Our Cleveland results are consistent with previous machine-learning studies that report substantial gains when moving from single-biomarker baselines to multivariable models on the UCI Cleveland benchmark.[11,12] Likewise, prior survival-prediction work in cardiology has shown that combining clinical variables can markedly improve mortality-risk stratification, aligning with our finding that multidimensional models outperform cholesterol-only approaches in the Faisalabad cohort. Because our goal was to compare cholesterol's role across two clinical contexts rather than exhaustively optimize algorithms, we limited model families to logistic regression and an MLP. Future work can extend this comparison to additional model classes (e.g., random forests, support vector machines, gradient boosting) using the same train/test protocol for fair benchmarking.

Finally, ROC curves summarize discrimination for both datasets and models (Figure 7). On Cleveland, LR AUC = 0.91 and MLP AUC = 0.88; on Faisalabad, LR AUC = 0.77 and MLP AUC = 1.00; models were evaluated on held-out test splits across five runs. Overall, results partly revise our initial expectation: cholesterol alone was more prognostic for mortality (Faisalabad) than diagnostic for incident disease (Cleveland). However, adding broader clinical information consistently improved performance for both endpoints. LR offers clinician-friendly interpretability via coefficients and ORs, while MLPs can capture complex, nonlinear relations and deliver higher accuracy for survival prediction. A dual-model strategy—transparent LR for screening and higher-capacity MLPs for prognosis—can balance interpretability and performance across clinical contexts.

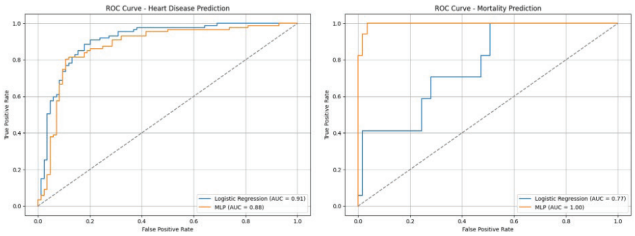


Figure 7: Receiver-Operating-Characteristic (ROC) curves comparing LR and MLP across both datasets. LR slightly outperformed MLP for incident disease in Cleveland (AUC 0.91 vs 0.88), whereas MLP dominated for mortality in Faisalabad (AUC 1.00 vs 0.77). This highlights a dual-model strategy: interpretable LR for screening and high-capacity MLP for prognosis.

Conclusion

Logistic regression and MLP models are effective machine-learning tools for predicting cardiovascular outcomes using clinical, demographic, and cholesterol data. To compare the predictive power of multidimensional feature sets with cholesterol alone, we used the Faisalabad Institute of Cardiology dataset for mortality and the UCI Cleveland Heart Disease dataset for incident heart disease. Our data show that cholesterol alone cannot reliably predict heart disease; in the Cleveland sample, it achieved a moderate accuracy of 53%. By adding age, sex, blood pressure, maximum heart rate, ST-segment depression, comorbidities, and lifestyle factors, logistic regression reached AUC = 0.91 and 85% accuracy.

In the Faisalabad cohort of previously diagnosed patients, cholesterol alone predicted mortality with 78% accuracy. Adding demographic and clinical data increased logistic-regression accuracy to 82% and MLP accuracy to 97%. Diabetes raised mortality risk 2.61×, while mortality increased by ~62% per unit of cholesterol; higher ECG ST-segment slope values were also associated with increased risk. These findings emphasize multidimensional risk models for cardiovascular disease diagnosis and prognosis. Logistic regression provides interpretability and clinical reasoning through odds ratios and coefficients, while MLP models improve survival-prediction accuracy and capture nonlinear interactions. This trade-off supports a two-model approach: transparent LR for initial screening and higher-capacity MLPs for individualized prognostication.

This study provides cardiovascular risk-stratification insights despite limitations in dataset size, single-center sam-

pling, and the absence of lifestyle detail and lipid subtypes. Future work should incorporate temporal modeling, additional biochemical and lifestyle markers, and multi-site cohorts. Overall, combining traditional biomarkers like cholesterol with modern computational methods can enable more accurate, resource-aware care and better patient outcomes.

Limitations and Future Work

This study provides cardiovascular risk-stratification insights despite limitations in dataset size, single-center sampling for the Faisalabad cohort, and the absence of lifestyle detail, lipid subfractions (e.g., LDL/HDL), and medication information. Additionally, a small number of cholesterol values recorded as 0 in the Cleveland dataset may reflect the absence in the source data and could introduce bias in cholesterol-only analyses. To avoid overgeneralization, the reported performance should be interpreted as a comparative benchmark within these cohorts rather than a universally generalizable clinical tool.

Future work should incorporate external validation on independent cohorts, stronger validation practices (e.g., nested cross-validation and calibration analysis), and richer biochemical and lifestyle markers. Where available, future analyses should replace total cholesterol with full lipid panels and evaluate sensitivity to handling of zero/placeholder values (e.g., treating them as missing and applying imputation or exclusion). Finally, additional model families (e.g., random forests, SVMs, gradient boosting) can be compared under the same evaluation protocol to assess whether performance improvements are robust across algorithms.

Acknowledgments

The author extends sincere gratitude to his mentors and teachers for their invaluable guidance throughout this research. The author also thanks Saint Francis High School, Mountain View, California, for providing academic support and resources that made this study possible.

References

1. Roth, G. A.; Mensah, G. A.; Johnson, C. O.; Addolorato, G.; Ammirati, E.; Baddour, L. M.; Barengo, N. C.; Beaton, A. Z.; Benjamin, E. J.; Benziger, C. P.; Bonny, A. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update from the GBD 2019 Study. *J. Am. Coll. Cardiol.* **2020**, *76* (25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010
2. Vourakis, M.; Mayer, G.; Rousseau, G. The Role of Gut Microbiota on Cholesterol Metabolism in Atherosclerosis. *Int. J. Mol. Sci.* **2021**, *22* (15), 8074. https://doi.org/10.3390/ijms22158074
3. Bentzon, J. F.; Otsuka, F.; Virmani, R.; Falk, E. Mechanisms of Plaque Formation and Rupture. *Circ. Res.* **2014**, *114* (12), 1852–1866. https://doi.org/10.1161/CIRCRESAHA.114.302721
4. Casula, M.; Colpani, O.; Xie, S.; Catapano, A. L.; Baragetti, A. HDL in Atherosclerotic Cardiovascular Disease: In Search of a Role. *Cells* **2021**, *10* (8), 1869. https://doi.org/10.3390/cells10081869
5. Mozaffarian, D.; Afshin, A.; Benowitz, N. L.; Bittner, V.; Daniels, S. R.; Franch, H. A.; Jacobs, D. R., Jr.; Kraus, W. E.; Kris-Etherton, P. M.; Krummel, D. A.; Popkin, B. M.; Whitsel, L. P.; Zakai, N. A. Population Approaches to Improve Diet, Physical Activity, and Smoking Habits: A Scientific Statement from the American Heart Association. *Circulation* **2012**, *126* (12), 1514–1563. https://doi.org/10.1161/CIR.0b013e318260a20b
6. Enkhmaa, B.; Berglund, L. Lipoprotein(a) and Effects of Diet: Time for Reassessment. *Nutrients* **2025**, *17* (10), 1714. https://doi.org/10.3390/nu17101714
7. Quispe, R.; Elshazly, M. B.; Zhao, D.; Toth, P. P.; Puri, R.; Virani, S. S.; Blumenthal, R. S.; Martin, S. S.; Jones, S. R.; Michos, E. D. Total Cholesterol/HDL-Cholesterol Ratio Discordance with LDL-Cholesterol and Non-HDL-Cholesterol and Incidence of Atherosclerotic Cardiovascular Disease in Primary Prevention: The ARIC Study. *Eur. J. Prev. Cardiol.* **2020**, 27 (15), 1597–1605. https://doi.org/10.1177/2047487319862401
8. Cushman, M.; Shay, C. M.; Howard, V. J.; Jiménez, M. C.; Lewey, J.; McSweeney, J. C.; Newby, L. K.; Poudel, R.; Reynolds, H. R.; Rexrode, K. M.; Sims, M. Ten-Year Differences in Women's Awareness Related to Coronary Heart Disease: Results of the 2019 American Heart Association National Survey. *Circulation* **2021**, *143* (7), e239–e248. https://doi.org/10.1161/CIR.0000000000000907
9. Kopčeková, J.; Holovičová, M.; Gažarová, M.; Mrázová, J.; Habánová, M.; Mečiarová, L.; Bronkowska, M. Association between Selected Dietary Habits and Lipid Profiles of Patients with Cardiovascular Disease. *Int. J. Environ. Res. Public Health* **2020**, *17* (20), 7605. https://doi.org/10.3390/ijerph17207605
10. Chicco, D.; Jurman, G. Machine Learning Can Predict Survival of Patients with Heart Failure from Serum Creatinine and Ejection Fraction Alone. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 16. https://doi.org/10.1186/s12911-020-1023-5
11. Yahaya, L.; Oye, N. D.; Adamu, A. Performance Analysis of Some Selected Machine Learning Algorithms on Heart Disease Prediction Using the Noble UCI Datasets. *Int. J. Eng. Appl. Sci. Technol.* **2020**, *5* (1), 36–46.
12. Varun, L. S.; Saivardhini, M.; Srivarsha, P.; Balaram, V. V. Predictive Analytics for Heart Disease Using Machine Learning. *Int. J. Adv. Res. Appl. Sci. Eng. Technol. (IJARASEM)* **2024**, *11* (6).
13. D'Agostino, R. B., Sr.; Vasan, R. S.; Pencina, M. J.; Wolf, P. A.; Cobain, M.; Massaro, J. M.; Kannel, W. B. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation* **2008**, *117* (6), 743–753. https://doi.org/10.1161/CIRCULATIONAHA.107.699579
14. Kontush, A. *Cholesterol, Lipoproteins, and Cardiovascular Health: Separating the Good (HDL), the Bad (LDL), and the Remnant; Wiley*: Hoboken, NJ, **2024**. https://doi.org/10.1002/9781394158393

Authors

Arth Bhardwaj is a research collaborator with the University of Illinois Urbana-Champaign and a Saint Francis High School student in Mountain View, California. His research integrates machine learning, huge language models, and cybersecurity systems to study AI safety, automation, and ethical deployment. Arth co-authored an IEEE threat intelligence conference article, demonstrating his technical expertise and interest in new security paradigms. He continues to work on durable, transparent, and secure AI-driven infrastructures for society.

DOI: 10.36838/IJHSR88.56

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Restoring p53 Function in Sarcomas Using MDM2 Inhibitors

Arwen Shah

Greenwood High School, No. 8-14, Sarjapur Road, Bangalore, Karnataka, 560035, India; arwenshah0@gmail.com

ABSTRACT: MDM2 is an E3 ubiquitin ligase and the principal negative regulator of the tumor suppressor gene TP53, maintaining cellular homeostasis through a tightly controlled feedback loop in which p53 transcriptionally activates *MDM2*, and MDM2, in turn, ubiquitinates p53 for degradation. In several sarcomas, particularly well-differentiated and dedifferentiated liposarcomas, *MDM2* amplification at chromosome 12q15 suppresses p53 activity despite retention of a wild-type TP53 gene. This makes the p53-binding pocket of MDM2 an attractive therapeutic target in these tumors. Over the past two decades, eight small-molecule MDM2 inhibitors have entered clinical trials, including RG7112, idasanutlin, SAR405838, HDM201, APG-115, navtemadlin, milademetan, and BI-907828, with seven advancing to later-stage evaluation. Although none have yet received regulatory approval, early-phase studies have demonstrated pharmacodynamic proof of concept, evidenced by p53 stabilization, induction of downstream targets such as p21, and tumor growth arrest in MDM2-amplified models and patients. This study compares the effectiveness of these inhibitors in the context of MDM2-amplified sarcomas, where excessive MDM2 expression suppresses wild-type TP53 activity. Pharmacologic blockade of the p53–MDM2 interaction can release p53 from inhibition, thereby restoring its tumor-suppressive functions and inducing cell-cycle arrest or apoptosis in cancer cells.

KEYWORDS: p53 Activation, MDM2 Inhibitor, Liposarcoma, Sarcoma Therapeutics, TP53 Regulation.

Introduction

The tumor suppressor protein p53, encoded by the TP53 gene, plays a crucial role in protecting the genome from damage. Often referred to as the 'guardian of the genome,' p53 halts the cell cycle and induces apoptosis (programmed cell death) or senescence (irreversible growth arrest) in response to DNA damage or cellular stress. In this way, p53 prevents the proliferation of potentially cancerous cells and helps maintain genomic stability.[1]

Under normal conditions, p53 activity is tightly regulated by MDM2 through an autoregulatory negative feedback loop.[2]

In several cancers, particularly well-differentiated and dedifferentiated liposarcomas (WDLPS/DDLPS), MDM2 is amplified at the 12q15 chromosomal locus, leading to excessive degradation of p53. Notably, many of these tumors retain wild-type (non-mutated) p53, which makes them particularly sensitive to MDM2 inhibition strategies.[3]

The therapeutic rationale for MDM2 inhibition is to disrupt the MDM2–p53 interaction and thereby restore p53's tumor-suppressive functions. This can reactivate p53 signaling in cancers with wild-type p53, resulting in cell cycle arrest or apoptosis.[4]

This review examines the therapeutic implications of MDM2 inhibition in sarcomas, with a focus on the evolution from Nutlin-3 to advanced clinical compounds, including RG7112, Idasanutlin, and Milademetan. Special attention is given to the underlying mechanisms of action, resistance, and combination strategies to improve treatment outcomes in liposarcoma patients.

Methods

Tumor suppressor gene (TP53):

The tumor suppressor gene TP53 encodes the protein p53, which regulates a multitude of genes to maintain homeostasis in the body. The transcription factor p53 (encoded by the TP53 gene) is often called the "guardian of the genome" because it plays a key role in maintaining genomic stability. TP53 regulates cellular stress responses through both transcriptional activation and repression of downstream target genes, depending on cellular context and damage signals.

Under normal (unstressed) conditions, p53 protein levels remain low due to continuous breakdown by the cell's protein recycling machinery. When cells experience stress such as DNA damage, low oxygen (hypoxia), or abnormal oncogene activation, p53 becomes stabilized and switched on. Once active, p53 turns on a set of target genes (Figure 1) that:

- Pause the cell cycle to allow DNA repair
- Trigger apoptosis (programmed cell death)
- Induce senescence (permanent cell-cycle arrest) if the damage is beyond repair

Canonical TP53 target genes include *CDKN1A* (p21), which mediates cell-cycle arrest; pro-apoptotic genes such as *BAX*, *PUMA*, and *NOXA*; and feedback regulators such as *MDM2*.

This protective process is so important that loss of p53 function happens in roughly half of all human cancers.[2] In tumors that have a normal (wild-type) *TP53* gene, p53 can still be inactivated indirectly by changes in its regulators.[5]

DOI: 10.36838/IJHSR88.61

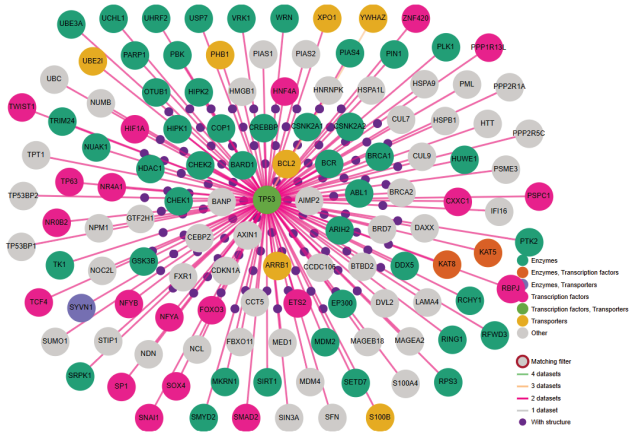


Figure 1: Interaction network of TP53 showing key transcriptional targets and regulatory partners involved in cell cycle control, apoptosis, and stress response (Human Protein Atlas). The above diagram was accessed from the Structure & Interaction resource of 'The Human Protein Atlas'. The Human Protein Atlas is a Global Core Biodata Resource in life sciences, based at SciLifeLab.

This network illustrates how TP53 integrates multiple cellular pathways, emphasizing its central role in maintaining genomic stability and its frequent disruption in cancer.

MDM2 as a negative regulator of TP53:

A primary negative regulator of p53 is MDM2 (Mouse Double Minute 2), an E3 ubiquitin ligase and proto-oncogene. MDM2 binds to the N-terminal transactivation domain of p53, blocking its transcriptional activity and adding ubiquitin tags that mark p53 for proteasomal degradation. MDM2 also transports p53 from the nucleus to the cytoplasm, promoting its degradation. This continuous surveillance maintains low basal p53 levels in unstressed cells. MDM2 itself is a transcriptional target of p53, creating an autoregulatory negative feedback loop as indicated in Figure 2: p53 activation induces MDM2 expression, and MDM2 then degrades p53 to restore homeostasis. This feedback ensures p53 is restrained under normal growth, but can rise rapidly under stress.[6] Under DNA damage or oncogenic stress, activation of ATM/ATR and downstream checkpoint kinases leads to phosphorylation of TP53 and MDM2, reducing their binding affinity and preventing ubiquitin-mediated degradation of p53.

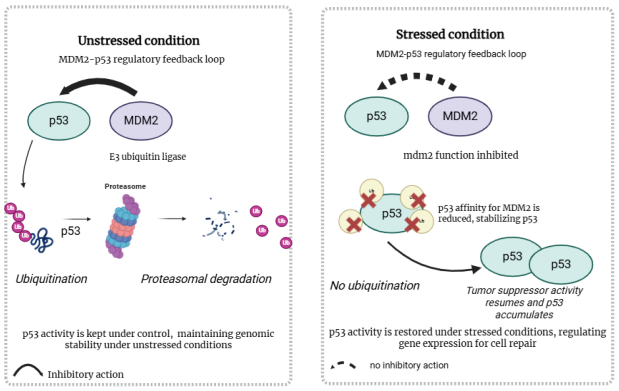


Figure 2a - Figure 2b: Regulation of p53 by MDM2 under homeostasis and stressed conditions. Under normal conditions, MDM2 ubiquitinates p53, maintaining low basal levels and preventing unnecessary cell-cycle arrest. Under stress, MDM2 inhibition stabilizes p53, enabling a rapid tumor-suppressive response through cell-cycle arrest or apoptosis.

Role of MDM2 in sarcomas:

MDM2 is a protein that plays a significant role in several types of sarcomas, which are malignant tumors that originate from mesenchymal tissues such as bone, fat, muscle, and cartilage.[7] These tumors include well-differentiated and dedifferentiated liposarcoma, a subtype arising from adipocytic (fat) tissue that frequently shows high-level amplification.[8] MDM2 amplification is a defining hallmark of such sarcomas, and its interaction with the tumor suppressor p53 has emerged as an actionable therapeutic target.[9]

In certain sarcomas, particularly well-differentiated and dedifferentiated liposarcomas (WDLPS/DDLPS), the regulatory axis between MDM2 and p53 is hijacked by MDM2 gene amplification at chromosome 12q13–158 (Figure 3). A majority of these tumors retain wild-type TP53, meaning that p53's inherent tumor-suppressive function is not lost to mutation. Instead, p53 is rendered inactive upstream by MDM2 overactivity. Alternative splicing of *MDM2* has been reported in high–mutational-burden sarcomas, though current inhibitors largely target conserved N-terminal binding sites; data on variant-specific drug sensitivity remain limited.

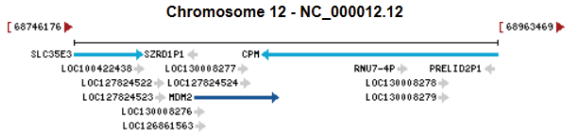


Figure 3: Location of the human *MDM2* gene on chromosome 12 (GRCh38/hg38 assembly) and corresponding nucleotide coordinates. This genomic map highlights the placement of *MDM2* within the 12q15 region, a locus frequently amplified in sarcomas, particularly well-differentiated and dedifferentiated liposarcomas. Such amplification drives oncogenesis by increasing MDM2 expression, which in turn suppresses p53 activity and promotes tumor growth. Source: NCBI Gene [Gene ID: 4193], National Center for Biotechnology Information. Available at: https://www.ncbi.nlm.nih.gov/gene/4193/

For patients with advanced soft tissue sarcomas, treatment options remain limited, particularly once the disease recurs or metastasizes. Standard chemotherapy often produces short-lived responses, and durable disease control is uncommon. This therapeutic gap has driven interest in molecularly targeted approaches, especially in sarcoma subtypes where well-defined genetic alterations, such as *MDM2* amplification in TP53–wild-type tumors, provide a clear biological rationale for intervention.

■ Results

Clinical Development of MDM2 Inhibitors:

Building on the biological framework described above, MDM2 inhibitors have been developed to therapeutically exploit p53 suppression in MDM2-amplified sarcomas.

Well-differentiated and dedifferentiated liposarcomas are among the sarcoma subtypes most closely linked to MDM2 dysregulation. These tumors frequently harbor high-level MDM2 amplification while retaining wild-type TP53, creating a setting in which p53 function is suppressed rather than genetically lost. As a result, pharmacologic disruption of the MDM2–p53 interaction offers a direct means of restoring tumor-suppressive signaling. This genetic alignment has made

liposarcoma a primary clinical focus for the development and evaluation of MDM2 inhibitors.

Following the proof-of-concept with Nutlin-3, successive generations of small-molecule inhibitors have been developed and evaluated in preclinical and clinical settings. These key compounds and their effects *in vivo* and *in vitro* are reviewed in Table 1 below.

Across MDM2 inhibitors evaluated in sarcoma, efficacy reporting has been heterogeneous, with formal survival metrics such as median PFS, OS, and hazard ratios available only for select agents, most notably milademetan.

Table 1: The list of MDM2 inhibitors and their effects.

Compound	Unique Feature or Strategy	Key Preclinical Findings	Clinical Findings / Status	Reference
Nutlin-3	First-in-class; revealed proteotoxic stress and combination opportunities	Strong p53 activation; synergizes with chemotherapy and ferroptosis inducers	Proof-of-concept only; not clinically developed	Vassilev *et al.*[9]
RG7112	First clinical-stage oral Nutlin-based inhibitor	Stabilizes p53; tumor growth inhibition in MDM2-amplified models	Limited to disease stabilization; high hematologic toxicity	Ray-Coquard[3]
Idasanutlin (RG7388)	Improved potency; synergy with kinase and proteasome inhibitors	Potent p53 reactivation and tumor regression	AML Phase III failure; limited sarcoma data	Khurana *et al.*[10]
Milademetan (RAIN-32)	Intermittent dosing (3/14) reduces hematologic toxicity	Potent p53 induction; tumor growth suppression	Phase III in liposarcoma showed comparable efficacy to chemotherapy with better safety	Gounder *et al.*[11]
ALRN-6924	Stapled peptide; dual MDM2/MDMX inhibition; less myelosuppression	Effective dual inhibition; overcomes MDMX-driven resistance	Disease control with minimal hematologic toxicity	Saleh *et al.*[12]
Navtemadlin (KRT-232)	Highly potent; radiosensitising properties	Enhances radiation-induced tumor cell death	Ongoing sarcoma+radiation combo trials; Phase III in myelofibrosis	Hanna *et al.*[13]
MI-77301 (SAR405838)	Spirooxindole scaffold; apoptosis inducer	Induces apoptosis in preclinical DDLPS models	Early clinical activity was limited; further development slowed	Bill *et al.*[14]
BI-907828	Long-acting oral spirooxindole analog; sustained p53 activation	Complete regressions in DDLPS patient-derived xenografts	Promising early clinical results in sarcoma	Boehringer Ingelheim,[15]
Siremadlin (HDM201)	Selective 2nd-gen; synergizes with DNA-damaging chemotherapy	Tumor regression with combination approaches	Encouraging early-phase data with intermittent dosing	Novartis Pharmaceuticals[16]
APG-115	Immunomodulatory effects; enhances PD-1 (programmed cell death protein 1) blockade synergy	Activates antitumor immunity	Early signs of durable disease control in solid tumors	Fang *et al.*[17]

This table highlights the evolution of MDM2 inhibitors from the prototype Nutlin-3 to advanced clinical candidates. While Nutlin-3 provided proof of the therapeutic principle by stabilizing p53 and revealing stress-related vulnerabilities, subsequent molecules such as RG7112, MI-77301, BI-907828, and siremadlin (HDM201) demonstrated improved potency and clinical translatability.

Nutlin-3 (p53-MDM2 inhibitor):

Nutlin-3 is a landmark small-molecule inhibitor from the cis-imidazoline chemical series.[9] By occupying the p53-binding site on MDM2, it releases and stabilizes p53, rapidly reactivating its tumor suppressor functions (Figure 4a). At the sequence level, the specific amino acid residues involved in Nutlin binding can be visualized through Figure 4b.

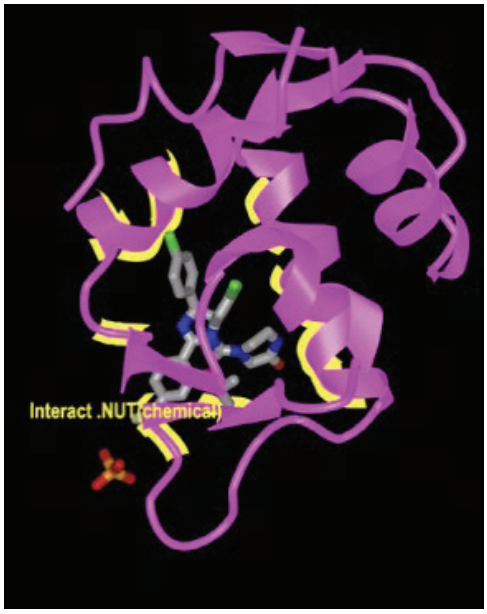


Figure 4a: Crystal structure of MDM2 bound to Nutlin-3a (PDB ID: 4HG7). The MDM2 protein is shown in ribbon format (magenta), and the Nutlin molecule is shown interacting at 11 residues, highlighted in yellow.

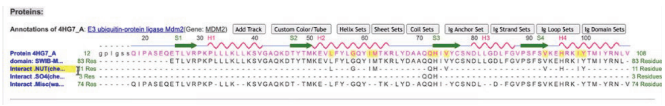


Figure 4b: Protein sequence annotation of MDM2 (UniProt: Q00987), displaying the structural features and positions of Nutlin-interacting residues. Source: These figures were generated using iCn3D ("I see in 3D"), a WebGL-based viewer developed by NCBI for interactive visualization of three-dimensional macromolecular structures and their chemical interactions.

In sarcoma models with wild-type TP53 and MDM2 overexpression, including Ewing sarcoma and osteosarcoma cell lines, Nutlin-3 strongly stimulated p53-dependent gene expression (notably p21and PUMA), resulting in cell-cycle arrest and apoptosis.[19] Significantly, Nutlin-3 induced robust tumor regressions in sarcoma xenograft models.[20] Beyond reactivating canonical p53 programs, Nutlin-3 also exposes new cellular vulnerabilities that can be exploited therapeutically:

Proteotoxic stress: Nutlin-3 treatment was shown to increase the production and breakdown of proteins inside liposarcoma cells, a process involving the proteasome.[21] This cellular com-

DOI: 10.36838/IJHSR88.61

plex degrades unwanted or damaged proteins. This surge in protein turnover creates proteotoxic stress; the cell accumulates misfolded proteins that overwhelm its quality-control systems. This stress activates a pathway involving ATF4 and CHOP, two transcription factors normally turned on during endoplasmic reticulum stress (a cellular alarm triggered by protein overload). When researchers knocked out these genes using CRISPR, Nutlin-3-induced cell death was reduced, proving that this stress pathway contributes to Nutlin's cytotoxicity. A genome-wide CRISPR screen also highlighted PSMD9, a component of the proteasome, as a key factor in Nutlin sensitivity; cells lacking PSMD9 were resistant to Nutlin-3, indicating that an active proteasome is necessary for the drug's full effect.

These findings led to rational combination strategies. Pairing Nutlin-3 with carfilzomib (a proteasome inhibitor) amplified cell death: by further weakening the proteasome, carfilzomib heightened the stress caused by Nutlin-3, boosting pro-apoptotic signals like CHOP and NOXA and causing dramatic apoptosis in sarcoma models.

Ferroptosis sensitivity: In preclinical model systems, Nutlin-3 has been shown to sensitize certain p53–wild-type tumor cell lines to ferroptosis, an iron-dependent form of cell death caused by lipid peroxidation. p53 activation by Nutlin-3 represses the cystine transporter SLC7A11, lowering the cell's antioxidant defenses and making it more vulnerable to ferroptosis inducers such as erastin. Pre-treating cells with Nutlin-3 followed by erastin led to markedly increased lipid ROS and cell death compared to either drug alone.[22]

DNA damage and extrinsic apoptosis pathways: Nutlin-3 also synergizes with chemotherapies such as doxorubicin. In breast cancer models, Nutlin-3 enhanced doxorubicin-induced apoptosis by boosting p53 activity.[23] Interestingly, Nutlin-3 can also engage the related p73 tumor suppressor pathway: in some p53-mutant cancers, Nutlin-3 triggers apoptosis via p73 activation.[24] This suggests that MDM2 inhibitors may retain partial activity even when p53 is mutated, expanding their potential therapeutic reach.

These vulnerabilities are beneficial in therapy design: they reveal stress points that combination drugs (e.g., proteasome inhibitors or ferroptosis inducers) can exploit to achieve stronger tumor killing than Nutlin-3 alone.

Despite strong preclinical activity, Nutlin-3 was unsuitable for clinical use due to its poor pharmacokinetic properties: high doses for effect, low oral bioavailability, rapidly cleared from the body.[25] These issues prompted the development of next-generation MDM2 inhibitors (e.g., RG7112, idasanutlin, milademetan) with improved drug-like properties.

RG7112:

RG7112 was the first clinical-stage, orally bioavailable MDM2 antagonist to enter human trials, derived directly from the Nutlin scaffold but optimized for improved binding and solubility. In preclinical studies, RG7112 bound to MDM2 with nanomolar affinity, reactivated p53 signaling in MDM2-amplified models, and induced tumor growth inhibition.[3]

A landmark Phase I trial tested RG7112 as a neoadjuvant therapy in patients with MDM2-amplified well- or dedifferentiated liposarcoma (WDLPS/DDLPS).[3] It clearly demonstrated increased intratumoral expression of p53 and p21, as well as reduced tumor proliferation. Most patients experienced disease stabilization, meaning tumor growth was halted rather than reversed. This study provided pharmacodynamic proof that p53 can be reactivated in MDM2-amplified sarcomas.

However, clinical utility was hindered by frequent hematologic side effects (notably thrombocytopenia and neutropenia) that necessitated dose interruptions. High-dose requirements and variable oral absorption also limited the long-term use of RG7112, so more potent and better-tolerated compounds succeeded it.

Idasanutlin:

Idasanutlin (RG7388) is a more potent, pyrrolidine-core MDM2 inhibitor with high oral bioavailability. Unlike Nutlin-derived molecules, idasanutlin features a pyrrolidine core that improves its fit within the MDM2 binding pocket and enhances its drug-like properties.[26] This structural innovation results in ~10-fold greater potency than RG7112 in cellular assays, with a reported MDM2 binding IC_{50} of ~6 nM.[26] (IC_{50} is the concentration of drug needed to inhibit 50% of its target activity; single-digit nanomolar values indicate extremely high potency, whereas earlier Nutlin compounds typically had IC_{50} values in the tens of nanomolar range.)

Idasanutlin also exhibits markedly better oral bioavailability, allowing for shorter and more convenient dosing regimens. Most clinical protocols use five consecutive days of treatment in a 28-day cycle, rather than the prolonged schedules required for RG7112.

Notably, idasanutlin demonstrated favorable combinations with proteasome inhibitors and FGFR/CDK4 inhibitors (reflecting key resistance escape mechanisms in liposarcoma). Clinically, most data are from AML trials (~25% CR with cytarabine), but early-phase sarcoma data show stable disease and occasional regressions. Hematologic side effects persisted but were less frequent/severe with intermittent dosing.[27]

One example is pairing idasanutlin with carfilzomib, a drug that blocks the proteasome (the cell's protein-degrading system). This combination created high levels of proteotoxic stress and triggered much stronger apoptosis than either drug alone. Another promising approach is combining idasanutlin with inhibitors of the fibroblast growth factor receptor (FGFR) pathway. FGFR signaling is often upregulated in DDLPS, providing tumor cells with an alternate survival route even when p53 is reactivated. In preclinical studies, combining idasanutlin with the FGFR inhibitor erdafitinib suppressed proliferation and induced apoptosis far more effectively than either drug alone, often achieving tumor regressions.[28] This synergy likely works as blocking FGFR removes one of the key pathways tumor cells use to resist p53-mediated stress.

Milademetan:

Milademetan is a third-generation oral MDM2 inhibitor specifically optimized for increased potency and tolerability in MDM2-amplified sarcomas. Its primary innovation is an intermittent dosing strategy (3 days on, 11 days off: "3/14"), developed to allow periods of p53 activation for antitumor effect while reducing sustained toxicity, particularly to the bone marrow. This schedule was shown in clinical trials to dramatically lower the risk and severity of side effects such as low platelets and neutropenia compared to continuous exposures seen with earlier MDM2 inhibitors.

In a Phase I study focusing on patients with advanced solid tumors, including a large dedifferentiated liposarcoma (DDLPS) expansion cohort, milademetan demonstrated meaningful clinical activity.[11] The overall disease control rate in DDLPS was as high as 58–62%, with a median progression-free survival nearing 7 months. Several patients experienced durable stable disease, and modest tumor shrinkages were observed, even if objective response rates remained low, consistent with the primarily cytostatic nature of p53 monotherapy in these tumors. Side effect profiles were manageable, with the majority of adverse events being mild and the intermittent schedule notably decreasing the frequency of more serious treatment-related cytopenias.

The subsequent Phase III MANTRA trial, comparing milademetan against trabectedin in advanced DDLPS, confirmed that oral intermittent MDM2 inhibition was as effective as standard chemotherapy in delaying tumor progression, but with fewer drug-related discontinuations and no treatment-related deaths. However, milademetan was not superior for progression-free survival or tumor response rates, so further development as a single agent for liposarcoma has been deprioritized.[29]

Although objective response rates were limited, the observed median progression-free survival of approximately 7 months in the DDLPS cohort exceeds historical outcomes with conventional chemotherapy, indicating that MDM2 inhibition primarily confers clinically meaningful disease stabilization rather than tumor regression.

Other MDM2 Inhibitors and Future Directions:

Beyond Nutlin-3, RG7112, idasanutlin, and milademetan, several additional MDM2-targeted therapies and related strategies are under development. These compounds either refine the Nutlin-class mechanism or expand into dual inhibition of *MDM2* and its homolog *MDMX* (also known as MDM4).

Together, they represent the next wave of attempts to fully reactivate wild-type p53 in cancers where MDM2 overexpression drives tumorigenesis.

MI-77301 (SAR405838):

MI-77301 (SAR202) is a spirooxindole-based MDM2 inhibitor designed for increased potency and selectivity compared to earlier compounds. It achieves tumor regression by restoring the p53 transcriptional program, effectively inducing apoptosis in models of dedifferentiated liposarcoma (DDLPS) that retain wild-type TP53 and have MDM2 amplification. Its efficacy is dependent on the presence of intact p53, making patient selection critical. Despite promising preclinical results, clinical progression has been limited, and development slowed to focus on more advanced analogs.[30]

BI-907828:

BI-907828 is an optimized, long-acting oral derivative of MI-77301. It has induced complete tumor regressions in DDLPS patient-derived xenograft models and has demonstrated promising clinical activity in early-phase sarcoma trials. This compound represents an advanced candidate for combination therapies aimed at improving clinical outcomes in sarcomas harboring MDM2 amplification.[31]

Navtemadlin (KRT-232 / AMG-232):

Navtemadlin (KRT-232) is an oral MDM2 inhibitor originally developed by Amgen (as AMG-232)[18] and now advanced by Kartos Therapeutics. It is one of the most potent MDM2 antagonists reported to date, showing substantially higher activity in biochemical and cellular assays compared to RG7112 or idasanutlin. In preclinical studies, navtemadlin induced rapid p53 stabilization, cell-cycle arrest, and apoptosis across multiple MDM2-amplified tumor models. It is currently in a Phase III trial for myelofibrosis, where the therapeutic goal is to eradicate malignant hematopoietic progenitor cells through p53 reactivation selectively. Navtemadlin is also being studied in combination with radiation therapy for soft tissue sarcomas (NRG-DT001) to determine whether MDM2 inhibition can radiosensitise these tumors, paralleling the synergy seen with Nutlin-3 in preclinical systems.[32]

ALRN-6924 (Sulanemadlin):

Tumors often upregulate MDMX as an escape mechanism under selective pressure from Nutlin-class drugs. ALRN-6924 (sulanemadlin) is a stapled peptide that mimics the alpha-helical region of p53 that interacts with both MDM2 and MDMX, thereby displacing p53 from both inhibitors simultaneously. In a Phase I trial in wild-type p53 solid tumors, ALRN-6924 demonstrated favorable tolerability, with notably lower rates of thrombocytopenia than small-molecule MDM2 inhibitors. The trial reported a 59% disease control rate, including some complete and partial responses. Although ALRN-6924's development pivoted toward chemoprotection, exploiting its ability to activate p53 in healthy tissues to shield them from chemotoxicity, its dual-targeting concept remains highly relevant for overcoming resistance in MDM2-amplified cancers.[12]

Siremadlin (HDM201):

Siremadlin (HDM201) is a second-generation oral MDM2 inhibitor with high selectivity and potency, designed to minimise off-target effects. In MDM2-amplified sarcoma preclinical models, siremadlin activated p53 signalling, upregulated p21, and caused growth arrest or apoptosis. Its combination with DNA-damaging chemotherapy showed synergistic effects.

In Phase I/II studies, siremadlin produced early signs of activity in MDM2-amplified sarcomas using intermittent

DOI: 10.36838/IJHSR88.61

dosing schedules to manage cytopenias. It is being evaluated in both solid and hematologic malignancies.[33]

APG-115:

APG-115 is a selective oral MDM2 inhibitor that not only restores p53 activity but also exerts immunomodulatory effects. "Preclinical studies indicate that APG-115 can enhance antitumor immune responses, particularly in combination with PD-1 blockade. Trials in patients with solid tumors, including sarcomas, have shown tolerable safety and preliminary evidence of durable disease control, especially when used in combination with immunotherapy.[34]

Resistance to MDM2 inhibitors:

Drug resistance is a major clinical challenge for MDM2 inhibitor therapy in sarcoma. One predominant mechanism is the emergence of TP53 mutations during treatment: prolonged exposure to Nutlin-class inhibitors often leads to selection of new missense mutations within the p53 DNA-binding domain, thereby inactivating its transcriptional and apoptotic functions and rendering MDM2 inhibition ineffective.[9] Laboratory studies have repeatedly demonstrated the evolution of such p53 mutations following Nutlin-3 adaptation, highlighting the need either to deliver combination or time-limited therapies that prevent this evolutionary escape, or to target mutant p53 directly with agents like APR-246.[21]

A second escape route is *MDMX* (*MDM4*) upregulation: as Nutlin-class drugs only break the p53-MDM2 interaction, many tumors rapidly increase *MDMX* expression to continue suppressing p53 even when MDM2 is blocked.[12] Dual MDM2/MDMX antagonists, such as ALRN-6924, are now in trials specifically to circumvent this bypass.

Resistance can also arise from shifts in the apoptotic threshold. As p53-driven apoptosis increases, tumor cells respond by upregulating anti-apoptotic proteins such as BCL-2 and BCL-XL. Combining MDM2 inhibitors with BH3 mimetics (such as venetoclax or navitoclax) can block these survival pathways and restore apoptotic death.[14]

Oncogenic signaling crosstalk also facilitates therapeutic escape. In liposarcomas bearing CDK4 amplification or elevated FGFR activity, additional pathway activation makes tumor cells less reliant on the MDM2–p53 axis. Rational combinations with CDK4/6 inhibitors (e.g., palbociclib) or FGFR inhibitors (e.g., erdafitinib) have shown resensitization in preclinical models.

Finally, the dosing schedule is key: continuous MDM2 inhibitor exposure accelerates the selection of resistant clones, while intermittent dosing, such as the 3/14 milademetan protocol, preserves efficacy and delays resistance. These resistance mechanisms and their therapeutic implications are summarized schematically in Figure 5.

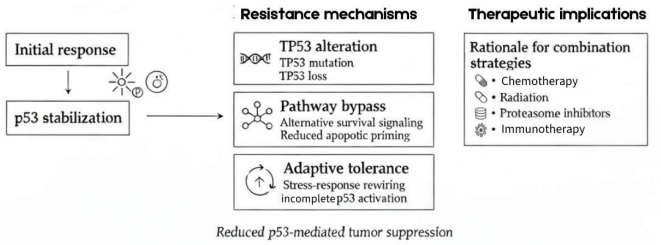


Figure 5: Conceptual overview of resistance mechanisms to MDM2 inhibition.

Evolving Strategies:

Emerging strategies for resistance include dual MDM2/MDMX inhibition by new-generation molecules, which block both arms of p53 repression and may preempt the most frequent bypass. Temporal sequencing and dynamic (primed or pulse) combination regimens are now utilized to heighten tumor vulnerability, for example, by first using an MDM2 inhibitor to synchronise sarcoma cells, then applying a targeted or cytotoxic second agent.

Beyond apoptosis, leveraging sensitization to extrinsic death via p53-driven upregulation of death receptors like TRAIL-R2, in combination with TRAIL agonists or immune modulators, is an active area of preclinical research.[3] Similarly, combining MDM2 inhibition with immunotherapy (e.g., PD-1/PD-L1 blockade) takes advantage of p53's role in fostering a pro-immunogenic tumor environment, and such approaches have entered clinical trials.

Finally, alternative cell death pathways, such as ferroptosis, are being targeted. Nutlin-class compounds sensitize cells to ferroptosis inducers (e.g., erastin), providing another tactic against apoptosis-resistant or heavily pretreated tumors. Adaptive, patient-specific therapy regimens shaped by real-time molecular monitoring are under development to further individualize and prolong MDM2 pathway responses. These are summarized in Table 2.

Table 2: Mechanisms of resistance to MDM2 inhibition and strategies to overcome them.

Cause of resistance	Mechanism	Impact on Therapy	Potential Strategies
New TP53 mutation	Mutations in p53's DNA-binding domain during treatment	Prevents p53 activation; the drug becomes ineffective	Shorten exposure; target mutant p53 (e.g., APR-246)
Upregulated MDMX expression	Overexpression of MDMX replaces MDM2's role	Still blocks p53 despite MDM2 drug	Use dual MDM2/MDMX inhibitors (e.g., ALRN-6924)
Higher anti-apoptotic proteins	Increased BCL-2/BCL-XL proteins block p53's apoptotic signal	Limits cell death despite p53 activation	Combine with pro-apoptotic drugs (e.g., venetoclax, navitoclax)
Alternate survival pathways	FGFR, CDK4, PI3K pathways activated	Cancer growth bypasses p53 dependency	Add pathway-specific inhibitors
Continuous dosing	Constant exposure favors resistant clones	Resistance appears faster	Use intermittent dosing schedules (e.g., milademetan 3/14 schedule)

This table emphasizes that newly acquired TP53 mutations are a critical resistance mechanism, rendering MDM2 inhibitors ineffective by abolishing p53's transcriptional activity. Importantly, it highlights how therapeutic strategies such as shortening exposure windows or targeting mutant p53 may help circumvent this resistance and extend the utility of MDM2 inhibitors.

Beyond biological resistance, several broader limitations have shaped the clinical trajectory of MDM2 inhibitors in sarcoma. Dose-limiting hematologic toxicities, particularly thrombocytopenia, have constrained continuous dosing strategies and limited therapeutic windows. In addition, objective tumor regressions remain uncommon with monotherapy, with clinical benefit often manifesting as disease stabilization rather than durable responses. Finally, variability in trial design and endpoints across studies complicates cross-agent comparison and has slowed regulatory progress in sarcoma indications.

Conclusion

The evolution from Nutlin-3 to second-generation idasanutlin and third-generation milademetan has firmly established the principle of pharmacologic p53 reactivation in cancer therapy. In MDM2-amplified sarcomas such as liposarcoma, where TP53 remains wild-type, these drugs exploit a unique vulnerability: the tumor's dependence on MDM2 overexpression to suppress p53.

Preclinical models have even demonstrated complete tumor regression with potent p53 reactivation, while clinical trials have shown prolonged disease stabilization in a subset of patients. However, consistent tumor shrinkage with monotherapy remains uncommon. Therapeutic challenges include the intrinsic p53–MDM2 (and MDMX) feedback loop, alternative tumor survival pathways, and on-target toxicities, particularly dose-limiting thrombocytopenia from p53 activation in normal tissues. Intermittent dosing, as with milademetan, has shown that efficacy and tolerability can be balanced, as confirmed by its Phase III trial, which demonstrated comparable activity but better safety than standard chemotherapy.

Future progress will rely on combination strategies that amplify p53-driven stress, such as pairing MDM2 inhibitors with proteasome inhibitors, ferroptosis inducers, TRAIL-receptor agonists, BH3 mimetics, or targeted kinase inhibitors. Dual MDM2/MDMX inhibitors may overcome resistance mediated by *MDMX* upregulation, and biomarker-guided patient selection, focusing on high MDM2 amplification, low MDMX expression, and intact p53, will be essential for maximizing benefit.

Next-generation agents like BI-907828 and navtemadlin promise greater potency, improved pharmacokinetics,[34] and potential synergy with immune checkpoint blockade, enabling p53 reactivation to work in concert with immune-mediated tumor clearance. With these refinements, the field is moving toward liberating the guardian of the genome into durable clinical benefit for sarcoma patients.

Acknowledgments

I would like to express my gratitude to my mentor, Dr. Lalita Limaye, for guiding me throughout my research project.

References

1. Wu, L.; Maki, C. G. *MDM2 Oncoprotein Functions*. In Madame Curie Bioscience Database [Internet], Landes Bioscience: Austin, TX; https://www.ncbi.nlm.nih.gov/books/NBK6130/.
2. Marei, H. E.; Althani, A.; Afifi, N.; Hasan, A.; Caceci, T.; Pozzoli, G.; Morrione, A.; Giordano, A.; Cenciarelli, C. p53 Signaling in Cancer Progression and Therapy. *Cancer Cell Int*. **21**, 703. https://doi.org/10.1186/s12935-021-02396-8.
3. Ray-Coquard, I.; Blay, J.-Y.; Italiano, A.; Le Cesne, A.; Penel, N.; Zhi, J.; Heil, F.; Rueger, R.; Graves, B.; Ding, M.; *et al*. Effect of the MDM2 Antagonist RG7112 on the p53 Pathway in Patients with MDM2-Amplified, Well-Differentiated or Dedifferentiated Liposarcoma. *Lancet Oncol*. **13** (11), 1133–1140.
4. Tisato, V.; Voltan, R.; Gonelli, A.; Secchiero, P.; Zauli, G. MDM2/X Inhibitors under Clinical Evaluation: Perspectives for the Management of Hematological Malignancies and Pediatric Cancer. *J. Hematol. Oncol*. **10**, 133. https://doi.org/10.1186/s13045-017-0500-5.
5. Oliner, J. D.; Saiki, A. Y.; Caenepeel, S. The Role of MDM2 Amplification and Overexpression in Tumorigenesis. *Cold Spring Harb. Perspect. Med*. **6**, a026336. https://doi.org/10.1101/cshperspect.a026336.
6. Yao, Y.; Zhang, Q.; Li, Z.; *et al*. MDM2: Current Research Status and Prospects of Tumor Treatment. *Cancer Cell Int*. **24**, 170. https://doi.org/10.1186/s12935-024-03356-8.
7. Sbaraglia, M.; Bellan, E.; Dei Tos, A. P. The WHO Classification of Soft Tissue Tumors: News and Perspectives. *Pathologica* **113** (2), 70–84. https://doi.org/10.32074/1591-951X-213.
8. Gambella, A.; Bertero, L.; Rondón-Lagos, M.; Verdun Di Cantogno, L.; Rangel, N.; Pitino, C.; *et al*. FISH Diagnostic Assessment of MDM2 Amplification in Liposarcoma: Potential Pitfalls and Troubleshooting Recommendations. *Int. J. Mol. Sci*. **24** (2), 1342. https://doi.org/10.3390/ijms24021342.
9. Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; *et al*. In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science* **303** (5659), 844–848. https://doi.org/10.1126/science.1092472.
10. Khurana, A.; Shafer, D. A. MDM2 Antagonists as a Novel Treatment Option for Acute Myeloid Leukemia: Perspectives on the Therapeutic Potential of Idasanutlin (RG7388). *Onco Targets Ther.* **12**, 2903–2910. https://doi.org/10.2147/OTT.S172315.
11. Gounder, M. M.; Rosen, L. S.; Patel, M. R.; *et al*. A First-in-Human Phase I Study of Milademetan (RAIN-32), an Oral MDM2 Inhibitor, in Patients with Advanced Liposarcoma, Solid Tumors, or Lymphomas. *J. Clin. Oncol.*
12. Saleh, M. N.; Patel, M. R.; Bauer, T. M.; Goel, S.; Falchook, G. S.; Shapiro, G. I.; *et al*. Phase 1 Trial of ALRN-6924, a Dual Inhibitor of MDMX and MDM2, in Patients with Solid Tumors and Lymphomas Bearing Wild-Type TP53. *Clin. Cancer Res*. **27** (19), 5236–5247. https://doi.org/10.1158/1078-0432.CCR-21-0715.
13. Hanna, G. J.; DeCaprio, J. A.; Mei, J. H. M.; McGreivy, J. S. An Open Label, Multicenter, Phase II Study of KRT-232, an Oral Small Molecule Inhibitor of MDM2, for the Treatment of Patients with Merkel Cell Carcinoma Who Have Failed Treatment with Anti-PD-1/L1 Immunotherapy. *J. Clin. Oncol*. **37** (15_suppl), TPS9602. https://doi.org/10.1200/JCO.2019.37.15_suppl.TPS9602.

DOI: 10.36838/IJHSR88.61

14. Bill, K. L. J.; Garnett, J.; Meaux, I.; Ma, X.; Creighton, C. J.; Bolshakov, S.; *et al*. SAR405838: A Novel and Potent Inhibitor of the MDM2:p53 Axis for the Treatment of Dedifferentiated Liposarcoma. *Clin. Cancer Res*. **22** (5), 1150–1160. https://doi.org/10.1158/1078-0432.CCR-15-1522.
15. Boehringer Ingelheim. This Study Aims to Find the Best Dose of BI 907828 (Brigimadlin) in Patients with Different Types of Advanced Cancer (Solid Tumors). *ClinicalTrials.gov*, Identifier NCT03449381. https://clinicaltrials.gov/study/NCT03449381.
16. Novartis Pharmaceuticals. Study to Determine and Evaluate a Safe and Tolerated Dose of HDM201 in Patients with Selected Advanced Tumors That Are TP53wt. *ClinicalTrials.gov*, Identifier NCT02143635. https://clinicaltrials.gov/study/NCT02143635.
17. Fang, D. D.; Tang, Q.; Kong, Y.; Wang, Q.; Gu, J.; Fang, X.; *et al*. MDM2 Inhibitor APG-115 Synergizes with PD-1 Blockade through Enhancing Antitumor Immunity in the Tumor Microenvironment. *J. Immunother. Cancer* **7** (1), 327. https://doi.org/10.1186/s40425-019-0750-6.
18. Gluck, W. L.; Gounder, M. M.; Frank, R.; Eskens, F.; Blay, J.-Y.; Cassier, P. A.; *et al*. Phase 1 Study of the MDM2 Inhibitor AMG 232 in Patients with Advanced p53 Wild-Type Solid Tumors or Multiple Myeloma. *Invest. New Drugs* **38** (3), 831–843. https://doi.org/10.1007/s10637-019-00840-1.
19. Arya, A. K.; El-Fert, A.; Devling, T.; Eccles, R. M.; Aslam, M. A.; Rubbi, C. P.; *et al*. Nutlin-3, the Small-Molecule Inhibitor of MDM2, Promotes Senescence and Radiosensitizes Laryngeal Carcinoma Cells Harboring Wild-Type p53. *Br. J. Cancer* **103** (2), 186–195. https://doi.org/10.1038/sj.bjc.6605739.
20. Sonnemann, J.; Palani, C. D.; Wittig, S.; Becker, S.; Eichhorn, F.; Voigt, A.; *et al*. Anticancer Effects of the p53 Activator Nutlin-3 in Ewing's Sarcoma Cells. *Eur. J. Cancer* **47** (9), 1432–1441. https://doi.org/10.1016/j.ejca.2011.01.015.
21. Ludwig, M. P.; Galbraith, M. D.; Eduthan, N. P.; Hill, A. A.; Clay, M. R.; Moreno Tellez, C.; *et al*. Proteasome Inhibition Sensitizes Liposarcoma to MDM2 Inhibition with Nutlin-3 by Activating the ATF4/CHOP Stress Response Pathway. *Cancer Res*. **83** (15), 2543–2556. https://doi.org/10.1158/0008-5472.CAN-22-3173.
22. He, W.; Shu, W.; Xue, L.; Wang, Y.; Chai, Y.; Wu, H.; *et al*. Synergistic Effect of Erastin Combined with Nutlin-3 on Vestibular Schwannoma Cells as p53 Modulates Erastin-Induced Ferroptosis Response. *J. Oncol*. **2022**, 7507857. https://doi.org/10.1155/2022/7507857.
23. Lee, D. M.; Kim, I. Y.; Seo, M. J.; Kwon, M. R.; Choi, K. S. Nutlin-3 Enhances the Bortezomib Sensitivity of p53-Defective Cancer Cells by Inducing Paraptosis. *Exp. Mol. Med*. **49** (8), e365. https://doi.org/10.1038/emm.2017.112.
24. Abrams, S. L.; Ruvolo, P. P.; Ruvolo, V. R.; Ligresti, G.; Martelli, A. M.; Cocco, L.; *et al*. Targeting Signaling and Apoptotic Pathways Involved in Chemotherapeutic Drug-Resistance of Hematopoietic Cells. *Oncotarget* **8** (45), 76525–76557. https://doi.org/10.18632/oncotarget.20408.
25. Urso, L.; Cavallari, I.; Silic-Benussi, M.; Biasini, L.; Zago, G.; Calabrese, F.; *et al*. Synergistic Targeting of Malignant Pleural Mesothelioma Cells by MDM2 Inhibitors and TRAIL Agonists. *Oncotarget* **8** (27), 44232–44241. https://doi.org/10.18632/oncotarget.17790.
26. Ding, Q.; Zhang, Z.; Liu, J. J.; Jiang, N.; Zhang, J.; Ross, T. M.; *et al*. Discovery of RG7388, a Potent and Selective p53-MDM2 Inhibitor in Clinical Development. *J. Med. Chem*. **56** (14), 5979–5983. https://doi.org/10.1021/jm400487c.
27. Phelps, D. A.; Bondra, K.; Seum, S.; Chronowski, C.; Leasure, J.; Kurmasheva, R. T.; *et al*. Inhibition of MDM2 by RG7388 Confers Hypersensitivity to X-Radiation in Xenograft Models of Childhood Sarcoma. *Pediatr. Blood Cancer* **62** (8), 1345–1352. https://doi.org/10.1002/pbc.25492.
28. Dadone-Montaudié, B.; Laroche-Clary, A.; Mongis, A.; Chamorey, E.; Di Mauro, I.; Chaire, V.; *et al*. Novel Therapeutic Insights in Dedifferentiated Liposarcoma: A Role for FGFR and MDM2 Dual Targeting. *Cancers (Basel)* **12** (10), 3058. https://doi.org/10.3390/cancers12103058.
29. Somaiah, N.; Tap, W. D. MDM2–p53 in liposarcoma: The need for targeted therapies with novel mechanisms of action. Cancer Treat. Rev. 2024, 122, 102668. https://doi.org/10.1016/j.ctrv.2023.102668
30. Wang, S.; Sun, W.; Zhao, Y.; McEachern, D.; Meaux, I.; Barrière, C.; *et al*. SAR405838: An Optimized Inhibitor of MDM2–p53 Interaction That Induces Complete and Durable Tumor Regression. *Cancer Res*. **74** (20), 5855–5865. https://doi.org/10.1158/0008-5472.CAN-14-0799.
31. Cornillie, J.; Wozniak, A.; Li, H.; Gebreyohannes, Y. K.; Wellens, J.; Hompes, D.; *et al*. Anti-Tumor Activity of the MDM2–TP53 Inhibitor BI-907828 in Dedifferentiated Liposarcoma Patient-Derived Xenograft Models Harboring MDM2 Amplification. *Clin. Transl. Oncol*. **22** (4), 546–554. https://doi.org/10.1007/s12094-019-02158-z.
32. Verstovsek, S.; Al-Ali, H. K.; Mascarenhas, J.; Perkins, A.; Vannucchi, A. M.; Mohan, S. R.; *et al*. BOREAS: A Global, Phase III Study of the MDM2 Inhibitor Navtemadlin (KRT-232) in Relapsed/Refractory Myelofibrosis. *Future Oncol*. https://doi.org/10.2217/fon-2022-0901.
33. Stein, E. M.; DeAngelo, D. J.; Chromik, J.; Chatterjee, M.; Bauer, S.; Lin, C. C.; *et al*. Results from a First-in-Human Phase I Study of Siremadlin (HDM201) in Patients with Advanced Wild-Type TP53 Solid Tumors and Acute Leukemia. *Clin. Cancer Res*. **28** (5), 870–881. https://doi.org/10.1158/1078-0432.CCR-21-1295.
34. Canon, J.; Osgood, T.; Olson, S. H.; Saiki, A. Y.; Robertson, R.; Yu, D.; *et al*. The MDM2 Inhibitor AMG 232 Demonstrates Robust Antitumor Efficacy and Potentiates the Activity of p53-Inducing Cytotoxic Agents. *Mol. Cancer Ther*. **14** (3), 649–658. https://doi.org/10.1158/1535-7163.MCT-14-0710.

Author

Arwen, a senior at Greenwood High International, India, studying the ISC curriculum, is passionate about the life sciences. She aspires to major in biotechnology and explore the complexities of living systems through research.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Optimal Thermoregulation Strategies for Building Roofs: A Multifaceted Approach Integrating Experimental Analysis and Finite Element Modeling

Ananya Chakravarthi

Plano East Senior High School, 3000 Los Rios Boulevard, Plano, Texas, 75094, USA; ay.chakravarthi@gmail.com

ABSTRACT: This study investigated building roof heating and cooling mechanisms through experimental testing and finite element modeling. We hypothesized that an optimum operating point exists for minimizing heat absorption while maximizing heat convection and radiation. Experiments were conducted on a model home with various roofing materials and coatings in a controlled environment. 3-D Finite Element Modeling simulations were performed to study different heating mechanisms. Results showed that impervious layers blocking convective cooling were less efficient than expected. A porous cover on the roof offered the best performance. Simulations highlighted the importance of radiative and convective cooling mechanisms in preventing excessive heating. Based on these findings, we propose a smart roof design to conserve energy utilization. This research combines experimental data from a smaller-scale prototype under controlled conditions with simulation to understand the problem of designing season-agnostic smart roofs, potentially leading to innovative technological improvements in building energy efficiency.

KEYWORDS: Energy: Physical, Sustainable Design, Heat Transfer Optimization, Finite Element Modeling (FEM).

Introduction

In countries such as the United States, ~40% of the total energy consumption is in buildings, with ~50% of residential housing energy used for heating and cooling to maintain a desirable indoor temperature (~22°C).[1] The choice of roofing material significantly impacts building heat retention and energy efficiency. As global temperatures rise and energy costs increase, optimizing building thermal performance has become a critical area of research. Gaining a physics-based understanding of heat transfer into and out of a home can help design and implement intelligent roofing systems that adapt to changing environmental conditions.

Current research in this field often focuses on individual aspects of heat transfer or specific materials, with much attention given to increasing/decreasing radiative cooling. Radiative cooling refers to the process by which surfaces emit thermal radiation to dissipate heat. In building applications, this allows roofs to cool below ambient air temperature, even during daytime. Past research on daytime radiative cooling, while successfully reducing cooling energy consumption, typically used materials with fixed, cooling-optimized properties.[1] These materials efficiently emit thermal radiation even when the surface temperature of the surface is lower than desired, such as at night or in the winter.[1] This unwanted thermal radiative cooling will increase the energy consumption for heating and may offset the cooling energy saved in hot hours or seasons.[1]

The research community acknowledges this issue well.[1,2] To cut the heating penalty from overcooling, techniques have been attempted to switch off thermal radiative cooling at low temperatures.[1] Although effective in switching, these techniques typically require either additional energy input or external activation, and in some cases, complex mechanical moving parts achieve switching.[1] Examples include temperature-controlled phase change structure,[3] deployable radiators with loop heat pipe, and 4 switchable cavitation coatings.[5] In an example of active control, Zhao *et al*. describe forming a bifunctional layer using a silicone film on a carbon block film that can be switched from one state that absorbs radiation to another state that allows radiative cooling.[5] On the other hand, a passive solution using vanadium oxide called temperature-adaptive radiative coating has been proposed that does not require active intervention.[1]

Despite this progress, such technology is still nascent and many years from commercial adoption, presumably requiring many years of development and testing. In addition, these solutions only consider minimizing radiative cooling and do not consider controlling conductive cooling, which is an equally important mechanism of heat transfer.

Thus, a significant gap exists in understanding how various heat transfer mechanisms interact in real-world conditions. While finite element modeling has been extensively applied to building thermal analysis, including roof heat transfer optimization and envelope performance studies, these approaches typically examine static material properties or focus on individual heat transfer mechanisms rather than investigating adaptive systems that can dynamically respond to seasonal variations. While it is nearly impossible to systematically study a real home's energy usage as a function of the external environment due to large variations in outside weather, a model home with a lower heat capacity can be studied within a constrained environment, providing valuable feedback.

This study aims to bridge this gap by combining experimental analysis with finite element modeling to investigate the complex interplay of conduction, convection, and radia-

DOI: 10.36838/IJHSR88.69

tion in roofing systems. Using a scaled model and controlled environment, we can isolate and analyze these heat transfer mechanisms more precisely than possible in a full-scale building. Our research specifically focuses on exploring the potential of porous materials and adaptive systems in roofing, hypothesizing that an optimal balance exists for a specific use case. This approach differs from traditional methods that often prioritize insulation at the expense of beneficial heat transfer during cooler periods.

The goal of this work is twofold. First, we want to comprehensively understand the thermal behavior of various roofing materials and designs under controlled conditions. Second, we want to use this understanding to propose a smart roof design that can adapt to seasonal changes, minimizing energy consumption year-round.

By integrating experimental data with computer simulations, we aim to provide insights that can guide the development of more energy-efficient and climate-responsive building technologies. This innovative approach, combining small-scale experimental analysis with finite element modeling, offers a unique perspective on roofing system optimization. We are unaware of any other work that used a combination of experimental data performed on a smaller-scale prototype under a controlled environment and simulation to understand this problem of designing a season-agnostic smart roof.

The findings from this research can contribute significantly to sustainable architecture and energy conservation efforts in the built environment. They could potentially lead to the development of adaptive roofing systems that dramatically reduce energy consumption in buildings across various climates.

Background Information:

Conduction, convection, and radiation are the primary heat transfer mechanisms in building roofs. While these mechanisms are generally well understood in isolation, their complex interplay in real-world conditions presents significant challenges.

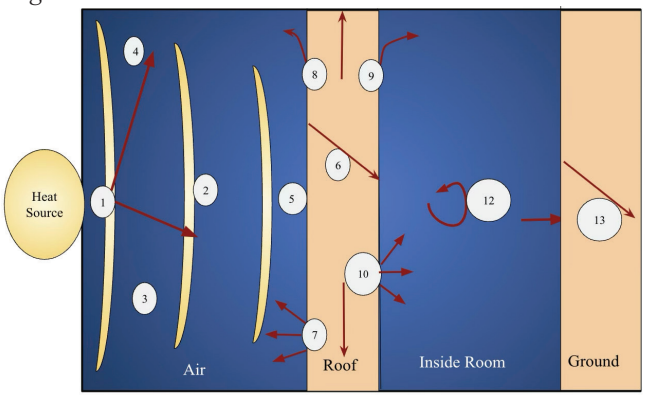


Figure 1: The various heat transfer mechanisms are at work when heat flows across the roof of a home. Straight arrows indicate conduction, curved arrows show convection, wavy lines represent radiation, and dashed lines show reflection. Numbered mechanisms correspond to detailed descriptions in the text.

1. Absorption by air of Electromagnetic Radiation (EMR) from a heat source: The air around the house absorbs a small fraction of the light from the heat source, typically contained in the infrared spectrum, increasing the ambient air temperature.

2. Conduction through air: Heat transfer occurs, especially in structures with a high surface-to-volume ratio, despite the air's poor thermal conductivity.

3. Convection through the air: Convection is the primary heat transfer phenomenon in fluids and is likely important. The hot air around the home can cause a convection current, with hot air rising away from the house and cold air moving toward the house.

4. Radiation through air: All objects emit radiation based on their temperature, including the air itself, which radiates heat in all directions.

5. Reflection / Absorption of heat at the roof: The roof absorbs any remaining electromagnetic radiation that is not absorbed by the air or reflected. The amount absorbed depends on the roof's material and color.

6. Conduction through the roof: Various layers of the roof conduct heat into the main body of the home. Traditional solutions focus on limiting this component, but this can also reduce beneficial heat gain during cooler periods, such as on a sunny winter day, as is typical in the Southwest US winters.

7. Radiation from the roof: The roof of the building radiates heat back into the atmosphere. It is typically the hottest part of the building.

8. Convection between the roof and air: As cold air from above contacts the warm roof, it absorbs heat and rises, contributing to heat transfer.

9. Convection between the roof and inside the house: Similar to the roof/air boundary, air currents within the home contribute to heat transfer.

10. Radiation in the home: The hot roof also radiates heat into the house's interior.

11. Conduction within the room: While likely playing a minor role, heat conducts through the air inside the house.

12. Convection within the room: As the room heats up, convection currents are established inside the house.

13. Heat transfer to the ground: Heat can transfer to the ground via the floor through conduction, radiation, and convection.

The complex interplay of these heat transfer mechanisms in building roofs presents challenges and opportunities for energy-efficient design. Our study explores how various roofing materials and designs, particularly porous materials and adaptive systems, can influence these heat transfer processes. By examining the balance between conduction, convection, and radiation, we aim to identify innovative roofing solutions that can effectively manage heat transfer across different seasons and climatic conditions, potentially leading to significant improvements in building energy efficiency.

■ Experimental Details: Method For Phase 1 Roof Experiment:

The steps below describe the process of collecting experimental data after building a model home, i.e., a smaller prototype with a lower thermal mass, which can be more easily

tested. The model home contains two separate units that are brought together: a base representative of the house's living space and a roof structure representative of the attic. The roof structure can be attached to the base to form a model home. The roof structure is designed to be a triangular prism typical of home roofs. The model home dimensions were as follows: a base unit measuring 11in by 6in by 4.5in (length × width × height), with roof structures forming triangular prisms. Critical thermal components included roof material thickness of 3 mm, wall thickness of 5 mm, and internal air volume of approximately 9 liters. Several different roof structures were fabricated and attached for testing as needed. The roof structures were identical in size and shape, but the only difference was what was on the outermost layer facing the heat lamps. Besides the black roof (control), these included roofs painted with lighter-colored and photochromic paint. Sometimes, a layer (aka sun shade) was attached to the roof instead of having a different coating. This layer could be an impervious fabric or meshes with varying degrees of porosity.

The following describes the experimental procedure for collecting the data.

The control roof (black spray-painted) is attached to the base unit. The model home is positioned on a table under two 75-watt halogen lamps above the roof surface, providing a combined 150 watts of heating power. While the exact heat flux reaching the model was not directly measured, this setup simulates realistic solar heating scenarios. The heating cycle is initiated. After turning on the lamps, temperatures within and on the walls are measured every 10 seconds for 3 minutes, every 30 seconds from 3 minutes to 10 minutes, every minute from 10 minutes to 25 minutes, and every 5 minutes from 25 minutes to 45 minutes (recorded in seconds). Prior testing showed that 45 minutes was sufficient to reach a steady temperature. After measuring the heating rate, the lamps are turned off, and temperature measurements during cooling begin at the same time intervals to initiate the cooling cycle. This process is repeated at least three times. Following the control experiment, the model home is removed, and a different roof configuration is attached, following Table 1, for the types of sunshades. For each roof, steps 2-5 are repeated to obtain a series of heating and cooling curves.

Table 1: Phase 1 variables.

Independent Variable	Constants	Dependent Variable
Type of sunshade: No mesh (CONTROL) 0% porous mesh 50% porous mesh 80% porous mesh Impermeable mesh Photochromatic vinyl Silver colored roof	Box (size, walls) Distance from the box to the bulb. Placement of the box on the table. Relative placements of thermometer, light sensor, and bulb. Time intervals.	Internal and external temperature in Celsius of the box

Phase 2 – Simulated Experiments:

In this research phase, we developed the structure of the model home in 3-D solid modeling software. We simulated some of the thermal mechanisms described above in 3-D Finite Element Modeling (FEM) software.

Background Information:

We explored various tools such as Energy2D and [6] Octave before deciding on the ANSYS OnScale platform.[7] OnScale Solve is a cloud-based engineering simulation software.[8] OnScale is a 3D Finite Element Modeling software that can solve static thermal equations. Because it was a cloud-based platform, it provided an easily accessible high-performance computing platform for performing mechanical, thermal, and fluid analyses. However, OnScale Solve requires input from a solid model (CAD).

OnShape is cloud-based 3D solid modeling software that generates solids.[9] The output from OnShape can be input into OnScale, allowing us to explore the designs.

Creating a 3D Solid Model:

In this project phase, we developed the structure of the model home, as shown in Figure 2, in a 3-D solid modeling software called OnShape. This involved the following basic steps.

1. Measurements: The dimensions of the experimental prototype from phase one were measured along with the design elements. All relevant design elements, including roof angles, wall thicknesses, and internal structures, were documented.
2. Model Development: A 3-D virtual prototype home with a hollow triangular prism roof with a solid side wall with the same dimensions as the experimental prototype was created in Ansys OnShape.
3. The 3-D virtual prototype home was placed on a virtual table, as in the experiments.

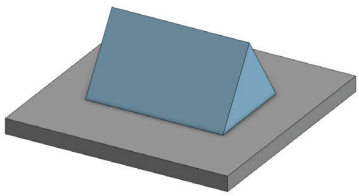


Figure 2: The 3-D Solid Model of the prototype gable roof. Figure 2 shows a prototype gable roof with the same dimensions as the experimental model created through OnShape.

Creating 3D FEM Model:

The steps below can be simulated to produce a thermal distribution, as shown in Figure 3. The solid model from OnShape was imported into OnScale, and various material and simulation parameters were set.

OnScale allows the materials to be selected. For our simulation, all the walls and roof of the virtual prototype home were chosen to be made of wood, while the table was selected to be made of plastic.

Next, the thermal simulation parameters and boundaries were set up in the Physics tab. The ambient temperature was selected to be 20 degrees C. A radiative boundary condition

DOI: 10.36838/IJHSR88.69

was added to all the surfaces by setting the emissivity to 0.8, which is representative of wood (which was changed in the special simulations below). In the experiments, the heat lamps faced only one roof surface, so a similar surface was selected to have a heat flux. The amount of heat flux was defined as 300 W/m2, which was the total heat power of the lamps combined divided by the area on the side of the roof facing the lamps. This heat flux of 300 W/m^2 represents approximately 30% of peak solar irradiance at Earth's surface (1000 W/m^2 at solar noon), simulating moderate daytime heating conditions rather than extreme summer solar loading.

All outside surfaces were included to have convective heat transfer. The convective heat transfer coefficient of air with no wind was selected as 12 W/m2C.

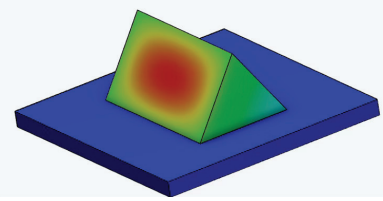


Figure 3: An example of thermal distribution in a simulation. Figure 3 shows the preliminary heat distribution on the prototype gable roof on the OnScale platform.

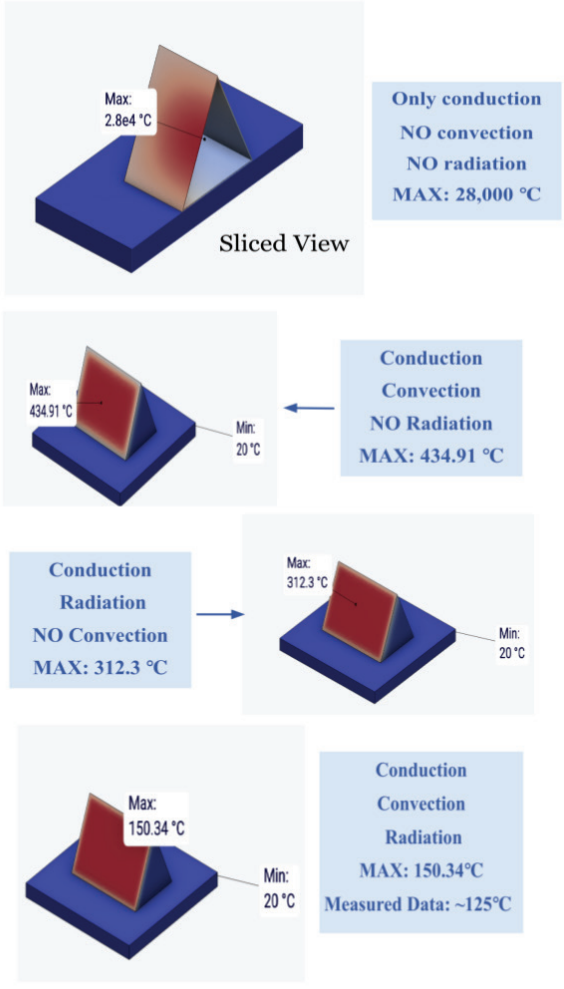


Figure 4: Four figures of the simulated conditions. Figure 4 shows the four trials tested on OnScale to understand the effect of heating mechanisms in a closed system.

The FEM simulations quantified the relative contribution of each heat transfer mechanism by systematically including or excluding specific physics. Figure 9 shows the temperature distributions for four simulation conditions, with the following maximum roof surface temperatures recorded:

Experiment (A) - Conduction Only: Maximum temperature of 245˚C, demonstrating that conduction alone produces unrealistically high temperatures due to a lack of heat dissipation mechanisms.

Experiment (B) - Conduction + Convection: Maximum temperature of 156˚C, showing that convective cooling reduces peak temperatures by 36% compared to conduction-only conditions.

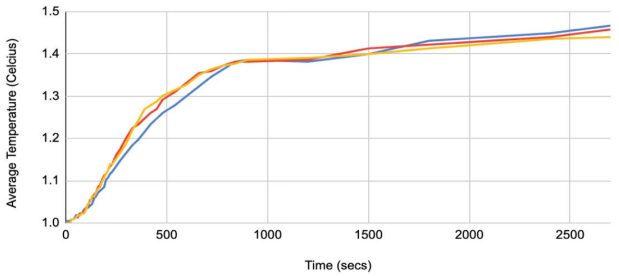
Experiment (C) - Conduction + Radiation: Maximum temperature of 78˚C, indicating that radiative cooling is more effective than convection for heat dissipation in this system.

Experiment (D) - All Mechanisms: Maximum temperature of 68˚C, closely matching experimental observations of approximately 65˚C and validating the simulation approach.

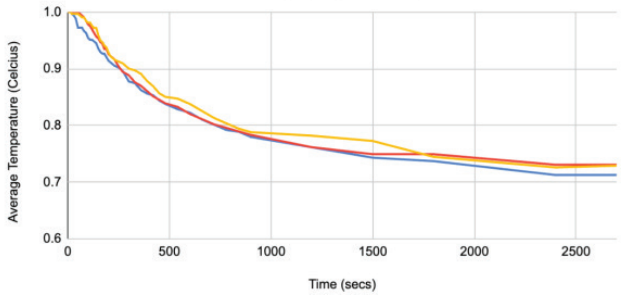
These results demonstrate that radiative cooling contributes most significantly to thermal regulation, while convective cooling provides important secondary benefits. The combined mechanisms are essential for realistic thermal modeling.

Results From Experimental Data

Figures 5A and 5B below show a representative normalized heating and cooling data plotted with time. During the early stages, there is a rapid increase in heating/cooling followed by a steady temperature.



(A)



(B)

Figure 5A: Heating curve plotting average (3 trials) temperatures of the black roof without mesh °C with time (s). Figure 5A shows rapid heating followed by reaching a quasi-steady state in which the temperature increase saturates.

Figure 5B: Cooling curve plotting average (3 trials) temperature of the black roof without mesh °C with time (s). Figure 5B shows rapid cooling followed by reaching a quasi-steady state in which the temperature decreases and saturates. followed by reaching a quasi-steady state in which the temperature increase saturates.

Results of Of Experimental Data

Figure 6 below compares the different types of roofing structures described above. Undoubtedly, the black roof causes the most heat to penetrate the prototype home, while the roof coated with silver paint blocks most of the heat. Both are understandable as the black painting absorbs the most heat (minimal reflection), and the silver paint (reflects the most heat).

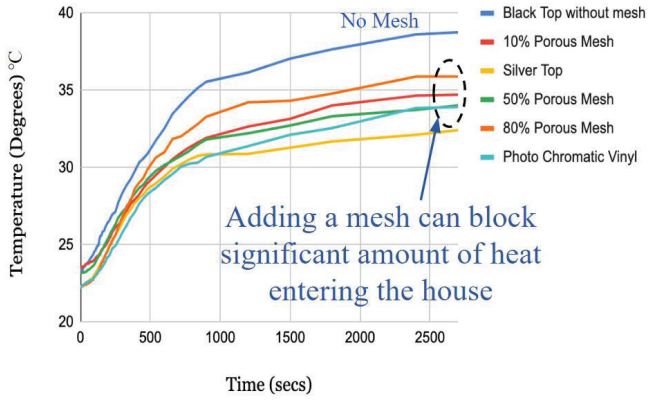


Figure 6: Heating curve comparing all sunshades by plotting time (s) with temperature °C. Figure 3 shows the difference in steady-state temperatures after rapid heating across 5 sunshades and a control (black top without mesh).

Figure 7 below compares the temperature reduction using a silvered roof (labeled as a colored roof) with a porous mesh. The silver-coated roof (silver paint coated over the black paint) is 20% better than the control black roof. The porous mesh below is about half as efficient as the silver-coated roof.

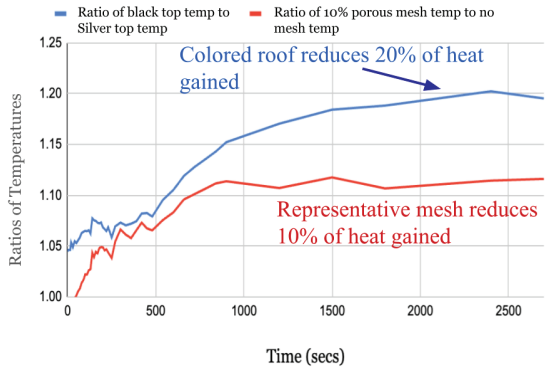


Figure 7: Ratios of temperature using color and mesh plotted with time (s). Figure 4 shows that the ratio of blacktop temperature to silver top temperature increases at a higher rate than the ratio of 10% porous mesh to no mesh.

Figure 8 below estimates the cooling rate for different roof structures with different porous layers. To estimate the cooling rate, the temperature difference for a short time window of 500s from the start of cooling was selected, as there is no difference at longer times. As the porosity of the mesh layer increases, the roof behaves almost like a roof without a mesh during cooling. In fact, at about 80% porosity, the mesh does not affect the cooling rate. While an impervious or highly non-porous layer is expected to form an insulative layer specifically by preventing convective heat transfer, the highly porous mesh layer does not limit convective heat transfer. Figure 8 reveals two distinct thermal behavior regimes separated by a critical transition around 50% porosity. Materials with low porosity demonstrate similar cooling rates, indicating that the mesh primarily functions as a radiation shield while limiting convective heat transfer. Beyond 50% porosity, the cooling rate increases dramatically, approaching the performance of unshielded surfaces as convective cooling dominates. This transition suggests that effective convective heat transfer requires a minimum threshold of pore connectivity, with isolated pores below 50% porosity unable to support meaningful airflow patterns.

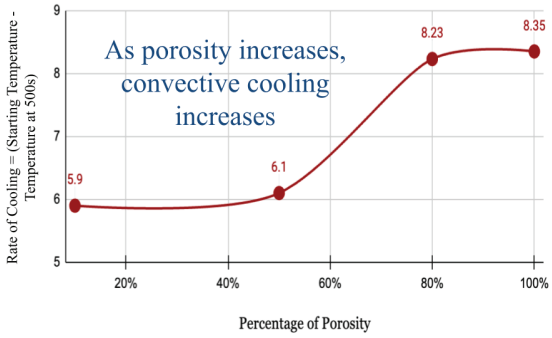


Figure 8: The cooling rate for different porous layers. Figure 5 shows that the rate of cooling, starting temperature minus temperature at 500 seconds, increases as the percentage of porosity increases.

Figure 9 below shows the steady-state temperature obtained with roofing structures with different porosities. Of course, the control with the black roof had the hottest room temperature. On the other hand, the prototype with the impervious vinyl was also quite hot at an average temperature of about 37 °C. The coldest structure is the structure with an intermediate porous layer attached to the roof.

This data is non-intuitive, and our simulation data explained below suggests that the following reasoning is likely correct. First, the cooling data clearly show that convective cooling of the roof surface plays an important part. Otherwise, the porous mesh will not cool faster than the less porous mesh (80% porous cools faster than 20% porous mesh). Based on the cooling data, it can be estimated that a highly porous mesh will allow convective cooling during heating compared to a less porous mesh. On the other hand, the highly porous mesh does not block the direct heat absorption on the roof's surface. For example, in an 80% porous mesh, 80% of the light energy will directly impact the roof surface, and much of that will be absorbed. On the other hand, in a 20% porous mesh, only 20% of the light energy will directly impact the roof surface. Thus, there is likely an optimum mesh porosity at which there is enough direct light energy is blocked while allowing some convective heat transfer. The experimental design was limited by available mesh materials, constraining us to discrete porosity values (0%, 50%, 80%, 100%). With greater flexibility in tailoring mesh porosities, more granular experimental data could precisely identify the optimal porosity range.

DOI: 10.36838/IJHSR88.69

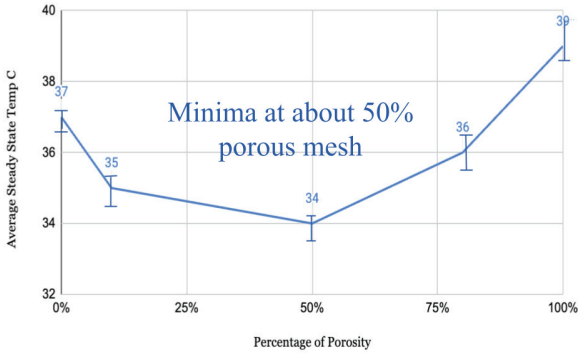


Figure 9: The average steady-state temperature °C of different porous meshes. Figure 6 shows a minimum of about 50% porous mesh when the average steady-state temperature of all variable porous meshes is compared.

Discussion Of Simulation Data

The maximum temperature value of the side facing the heat lamps is recorded in the simulations above in Figure 9. From our experimental data, I recorded the wall temperature for the control to be about 125 degrees. The final simulation (Experiment (D)) produced relatively comparable results. Note that the inside temperature within the model cannot be computed as the OnScale cannot simulate the airflow within the home caused by convection. Hence, the outside wall temperature was compared.

The simulation data quantifies each mechanism's contribution: convection reduces temperature by 89°C (245°C to 156°C), while radiation reduces temperature by 78°C (156°C to 78°C). However, convection is critical because without it, simulated temperatures (245°C) are 3.8 times higher than experimental values (~65°C), making the model unrealistic. This possibly explains why the minimally porous mesh layer attached to the roof provides the best result in the experiments. Unlike the impervious layer that blocks or eliminates most convective cooling, a porous mesh layer enables convective cooling while blocking direct radiation.

Phase III - Smart Roof Design:

Based on the knowledge gained from this work, a smart roof can be designed, as shown in Figure 10, to reduce energy consumption. This system can be a retractable porous fabric that can be mounted on the roof's side and receives the most sun (typically on the south side). This system could employ an active management system that could include different types of sensors. For example, sensors could roll open the fabric during the summer sun (blocking heat) and keep it retracted during the winter sun (allowing the sun's heat into the home to reduce heating bills). Unlike the various technologies under research, this technology is available and well-proven for reliability. For example, solar panels are installed on the roofs of many homes, and sunshades are installed on windows. Both technologies are currently available and can be easily repurposed.

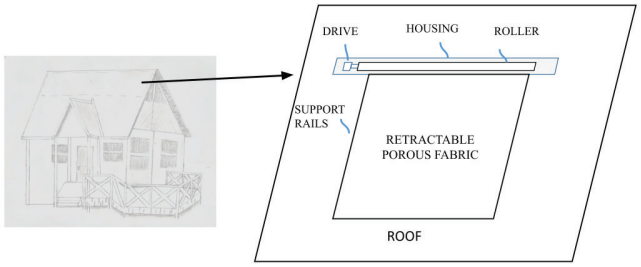


Figure 10: Figure 10 shows an active smart roof system that is relatively self-explanatory. A fabric is part of an electronic roll, similar to sunshades, but is supported on the house's roof.

Error Analysis of Experimental Data:

Before concluding with the data, the potential errors in our experimental data were reviewed. The observed results cannot be explained merely by experimental or statistical error. Possible sources of error include changes in ambient temperature and hysteresis between heating and cooling cycles. Due to the large amount of data collected during my school year, the experiment data was collected over a long period (about a month). During this time, the outside weather had some significant fluctuations, which caused the ambient temperature in my home, where all the experiments were performed, to fluctuate by a few degrees. However, since I normalized all the data, I do not think this could be a significant error source. Similarly, there was no significant hysteresis in the data. In addition, as shown in Table 2 below, the calculated standard error of the mean shown in the data is reasonably well-behaved.

Table 2: Standard error of the mean steady-state temperature (°C) for different porous mesh configurations based on three experimental trials. Table 2 shows that the standard error of the mean is low for all percentages of porous mesh.

Percentage of Transparency through Holes at steady state	Standard Error of the Mean
0%	0.26
10%	0.41
50%	0.40
80%	0.86
100%	1.1

Summary Of Experimental Data:

The experimental data clearly show that a porous sunshade is better than an impervious one. A roof with a 10-50% porous mesh is 10% better than one with an impervious mesh or no mesh. Although the silver-coated roof achieved superior thermal performance with a 20% improvement, the adaptive porous mesh system is recommended for practical implementation. The retractable mesh system offers dynamic control, allowing beneficial solar heating in winter while providing cooling in summer, making it more suitable for year-round energy optimization.

Conclusion

This study has successfully demonstrated the complex interplay of heat transfer mechanisms in building roofs and proposed an innovative approach to optimize thermal regulation. Through experimental analysis and finite element modeling, we have gained valuable insights into the thermal behavior of various roofing materials and designs. Our research revealed the effectiveness of porous materials in roof design, with a partially porous cover attached to the roof offering the best thermal performance. Specifically, a 10-50% porous mesh showed a 10% improvement in thermal regulation compared to impervious or non-mesh roofs. This finding underscores the importance of balancing heat blocking and dissipation, challenging the conventional approach of maximizing insulation at all times.

Both our experimental data and FEM simulations highlighted the significant impact of convective heat transfer in roof thermal regulation, a factor often underestimated in traditional roof designs. Based on these insights, we propose a smart roof design featuring a retractable porous fabric. This system could dynamically adapt to changing environmental conditions, optimizing energy efficiency across seasons.

These conclusions support our initial hypothesis and provide new insights that could spark a new generation of technological improvements in building energy efficiency. Our proposed smart roof system offers several advantages over existing solutions, including season-agnostic performance, targeted installation on high-heat absorption areas, potential for invisibility when retracted, and utilization of existing technologies for quicker adoption.

While our 3-D FEM modeling capabilities were limited, they provided crucial insights that complemented our accelerated testing experiments. This combined approach of small-scale prototyping and computational modeling offers a promising methodology for future research.

We envision collaborating with the North Texas Renewable Energy Group to discuss our findings and explore practical applications.[11] We also aim to partner with a roofing company to build and test a full-scale prototype of our smart roof design. Future work could involve enhancing our computational models to include more complex physics, such as improved convection modeling and time-dependent solutions, as well as investigating the potential of our adaptive roof system in various climatic conditions and building types.

In conclusion, this research contributes significantly to our understanding of roof thermal regulation and proposes a novel, practical solution for improving building energy efficiency. As global energy demands continue to rise and climate change presents increasing challenges, innovations in building design could play a crucial role in creating more sustainable and energy-efficient built environments. Our work paves the way for a new generation of adaptive, energy-efficient roofing systems that can significantly reduce energy consumption in buildings across diverse climates.

Acknowledgments

I would like to express my earnest gratitude to my parents for their endless support and guidance throughout this process, especially my father, without whom I would not be where I am today.

References

1. Tang, K.; Dong, K.; Li, J.; Gordon, M. P.; Reichertz, F. G.; Kim, H.; Rho, Y.; Wang, Q.; Lin, C.-Y.; Grigoropoulos, C. P.; Javey, A.; Urban, J. J.; Yao, J.; Levinson, R.; Wu, J. Temperature-Adaptive Radiative Coating for All-Season Household Thermal Regulation. Science 2021, 374 (6574), 1504–1509. https://doi.org/10.1126/science.abf7136.
2. Ulpiani, G.; Ranzi, G.; Kwok Wei, S.; Feng, J. On the Energy Modulation of Daytime Radiative Coolers: A Review on Infrared Emissivity Dynamic Switch against Overcooling. ResearchGate **2020**, 278–301. https://doi.org/10.1016/j.solener.2020.08.077.
3. Xia, Z.; Fang, Z.; Zhang, Z.; Shi, K.; Meng, Z. Easy Way to Achieve Self-Adaptive Cooling of Passive Radiative Materials. ACSPublications **2020**. https://doi.org/10.1021/acsami.0c05803.
4. Goncharov, K.; Orlov, A.; A. Tarabrin; M. Gottero; Perotto, V.; Tavera, S.; Zoppo, G. P. 1500 W Deployable Radiator with Loop Heat Pipe. SAE technical papers on CD-ROM/SAE technical paper series **2001**. https://doi.org/10.4271/2001-01-2194.
5. Zhao, H.; Sun, Q.; Zhou, J.; Deng, X.; Cui, J. Switchable Cavitation in Silicone Coatings for Energy-Saving Cooling and Heating. Advanced Materials **2020**, 32 (29), 2000870–2000870. https://doi.org/10.1002/adma.202000870.
6. Charles, X. Energy2D - Interactive Heat Transfer Simulations for Everyone. energy.concord.org. https://energy.concord.org/energy2d/.
7. Eaton, J. W. GNU Octave. octave.org. https://octave.org/.
8. Tomasetti, T. OnScale Solve - the Cloud Engineering Simulation Platform. OnScale. https://onscale.com/.
9. Onshape | Product Development Platform. www.onshape.com. https://www.onshape.com/en/.
10. Onshape - YouTube. www.youtube.com. https://www.youtube.com/channel/UCTvd5lUSLtTH8Qcd7Pl1nQg.
11. Optimal thermoregulation strategies for building roofs: a multifaceted approach integrating experimental analysis and finite element modeling, Invited presentation at North Texas Renewable Energy Group, https://www.ntreg.org/, August 10, 2024.

Author

Ananya Chakravarthi is a senior at Plano East Senior High School in Dallas, Texas. She aspires to study electrical engineering with a concentration in environmental science and is passionate about energy sustainability, Indian classical violin, and tennis.

DOI: 10.36838/IJHSR88.69

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

How Game Theory Affects the Video Games Industry

Richard L. Wan

Central Bucks High School East, 2804 Holicong Rd, Doylestown, PA 18902 USA; richardluwan2008@gmail.com

ABSTRACT: Despite the video gaming industry's ever-increasing engagement from both consumers and developers, it can be argued that it has somewhat reached a point of stagnancy. It is common to find complaints from consumers about its current overall state and how poorly it's performing. Why is the gaming industry, especially at the premium level, so unappealing for consumers? Why is it so easy to find complaints about how badly the field is performing despite its plentiful choices and apparent success? The modern video gaming industry is an extremely broad market that plays a major part in many people's lives, yet in recent times, there has been increasing dissatisfaction associated with the more premium levels of the medium, especially in the multiplayer genre, seemingly following very similar trends every time. If there is a pattern, then can it be explored such that a solution can be found that will satisfy customers and yet still appeal to a business's needs? The experiment consists of analyzing games based on user-based data. Then, dynamic modeling is used to draw conclusions from the data to support the claim that balancing between profit-based and customer-based thinking will ensure a healthier business environment for both parties.

KEYWORDS: Game Theory, Strategic Games, Video Games Industry, Profit Zone, Profit Driven, Community Driven.

Introduction

Every business's main goal is to be financially successful, and there are many ways to achieve it. However, while plenty of businesses are able to reach the end goal, there are a decent number of companies that fail to succeed for one reason or another. The most common reason is that a product doesn't appeal to enough people or drives them away. To verify this, the paper analyzes the live-service video games industry.

Firstly, to establish a baseline, we look at "The Profit Zone," a book written in 1997 by Adrian J. Slywotzky. The book provides insight into various uniquely applicable strategies in business that have persisted into modern-day, and, in our case, are especially applicable to the video games industry.[1]

The first of the strategies, the "product pyramid" model, provides options for different customers in different price ranges, from entry-level to premium. An example of this is Electronic Arts, a large video games publisher and developer, where there is a game for every price range in its value chain, most notably, a collection of free-to-play games at the "base" of the product pyramid and >$50 premium games at the "peak." This allows EA to be very accessible for a large number of age groups and spending profiles.[2]

The "switchboard" model not only creates and shares multiple products for consumers but also serves as a hub for other developers and businesses to spread their influence, allowing all parties involved to benefit. While there are many examples of a switchboard model in the video games industry, Valve's online video game store, Steam, stands out as one of the largest digital gaming distribution storefronts, providing a place for developers to sell their own products, both to consumers and other developers. This allows Steam to be very accessible to a large number of developers and to establish various connections between consumers and developers within its own network.[3]

The "total solution" model ensures that nearly any customer can be satisfied with one or more of the products along a value chain that a company provides. Epic Games is known for having a large online video game storefront – the Epic Games Store – for customers, a popular game-as-a-social-platform – Fortnite – for players, and a powerful, easily accessible game engine – Unreal Engine – for developers, all making up a thorough total solution model. This allows Epic Games to be a hub for all potential customers to find a satisfactory product in their value chain, regardless of what they are there for or what their background is.[4]

The "setting the standard" model can be applied when the company creates products that will revolutionize the market and encourage other companies to align with their standard; although a very difficult model to implement, it ensures that the market landscape is aligned with the company's interests, and will most often lead to success.[1] Nintendo, especially in the hardware market, has set the standard for handheld gaming consoles paired with high-quality games, historically with its Nintendo DS and Wii, and most recently with its Nintendo Switch in 2017, a handheld gaming console that is widely known for its high quality, longevity, and successful dive into an – at the time – lesser-used medium, setting the standard for an entire subcategory of console gaming to expand upon. This allows Nintendo to be a trendsetter that the rest of the industry looks to follow.[5]

In short, the "Profit Zone" strategies are used by a lot of modern businesses, and they, recently, have all culminated in a prevalent strategy: the "subscription" model, more commonly known as the "freemium" or "free-to-play" model. This ever-increasingly used business model, in its most prevalent form, encourages customers to play their game/use their product or service by providing all core features for free while monetizing

DOI: 10.36838/IJHSR88.76

additional content via cosmetics, expansions, and/or advertisements.[6]

To come out on top from the perspective of both parties, the product must establish a strong balance between customer and business to last in the long term.[1] To find said balance, companies need to use game theory strategically to analyze the landscape and successfully sell a product/service.[7]

Game theory is a field of mathematics that analyzes the various strategies in a game between two or more players, where each player has values assigned to each decision they can make; with game theory, predictions can be made on how the game will play out, and what decisions are most beneficial to each player. In business, game theory can be utilized to a great extent to analyze the actions made by both the consumer and competition, such that businesses can make well-informed decisions.[7]

Three major types of strategic games can be played and visualized between customers and businesses.

Coordination encourages – somewhat temporary – loyalty between both parties for greater benefit.

Collaboration is repeated cooperation that often results in a highly beneficial outcome for both parties when they rely upon and bounce off of each other, but it can also be easily undermined by one party deviating to gain a short-term advantage over the other, however minuscule it may be.

Competition occurs when cooperation isn't favorable for either side, and both sides have a clear intent to gain a significant, albeit possibly short-term, advantage over the other.[7]

The payoff matrices for each game are shown in Figure 1.

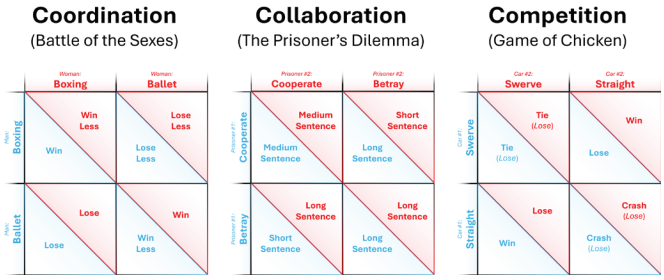


Figure 1: The payoff matrices for the "Battle of the Sexes", "Prisoner's Dilemma", and "Game of Chicken" games. The Battle of the Sexes favors cooperating, the Prisoner's Dilemma favors competition long-term and cooperation short-term, and the Game of Chicken is strictly competitive.

Competition between the customer and business occurs in the gaming industry due to business strategies being implemented, like "freemium," which encourages decisions that are unfavorable to the consumer, like paying for an expensive cosmetic that they might not have considered before.[8] Many patents from various companies, such as Activision Blizzard and Kabam, show exploitation of consumers through practices like FOMO – Fear of Missing Out – and giving consumers who pay an advantage over those who don't – Pay to Win. As a result, many games and companies that utilize such practices often gain a bad reputation with consumers.[9]

Gaming is a saturated market, and many strategies are exemplified by the businesses in the field. Strategic game theory and "profit zone" analyses are not only useful in the video game industry; however, they can also be projected onto other industries, allowing businesses in other fields to use the video game industry's strategies and appeal to customers with quality and support through collaboration, without losing potential profits in the long term.

The rest of the paper is organized as follows:

• Methods – a step-by-step walkthrough of how data is collected and analyzed.

• Results and Discussion – a presentation of the results found from the experiment and an explanation of how those results were reached.

• Conclusion – a summary of all the information that has been given until that point, and a proposition for how to further expand the topic.

Methods

To start, player data is collected from a selection of popular and/or notable live-service, multiplayer games in a multitude of genres, mainly collected from steamdb.info, a third-party website that contains data about all games on the online game playing and distribution platform Steam, by Valve Corporation.

Once the data is downloaded and organized in Excel, we focus on the user base over a game's lifespan and create a line graph from the exact data. The data from each time series is then categorized into 4 types of general graphs. Most of the games' time series align with the behavior of one of these graphs and, as such, can be categorized by their structure/pattern:

o Spike and Decline (250 days short-term)

o Spike and Increase (250 days short-term, 750 days long-term)

o Sustained Growth (>1000 days long-term)

o Growth and Decline (>1000 days long-term)

By categorizing games in this way, patterns may emerge between the games and their trends, which may align with our hypothesis.

Results

Tables 2 and 3 are the results of the data analysis. (τ) (days) stands for the time constant (the dynamics for the short and long term), and (Δ) (days) stands for the delay (the delay for the short- and long-term dynamics). The blue line is the real data, and the orange line is formed from the equations in Table 1.

Table 1: Equations for each graph type.

	Short Term	Long Term
Min to Max	Init = Min Value(t+1) = Value(t) + (Max – Value(t))/τ_short Value(t+1) = Value(t+1) * sigmoid	Max*(1±t/τ_long*sigmoid)
Max to Min	Init = Max Value(t+1) = Value(t) – (Value(t) – Min)/τ_short Value(t+1) = Value(t+1) * sigmoid	Min*(1±t/τ_long*sigmoid)
sigmoid	sigmoid = 1/(1+exp(-(t-Δ_short)/smoothing))	sigmoid = 1/(1+exp(-(t-Δ_long)/smoothing))

Table 2: Real and predicted graph data for each profit-driven game, along with the graph behavior and values produced from the equations.

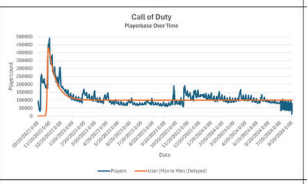
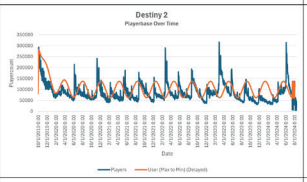
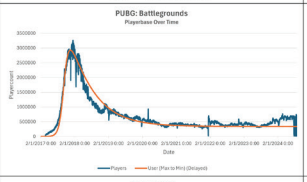
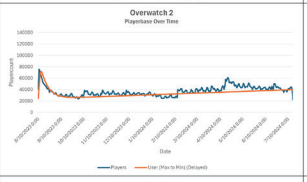
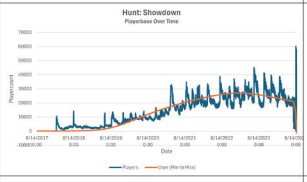
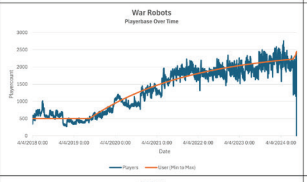
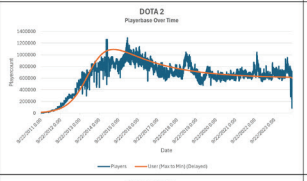
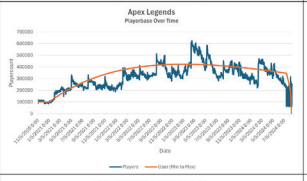
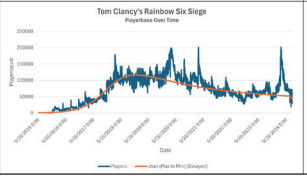
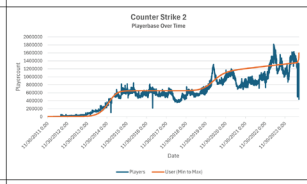
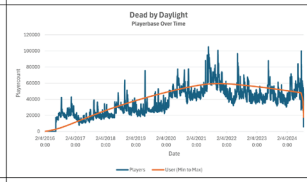
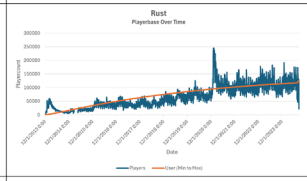
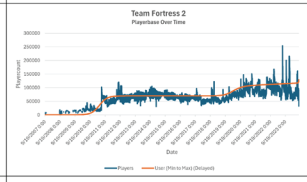
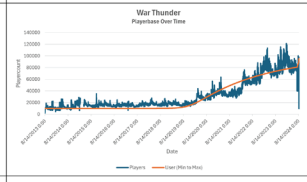
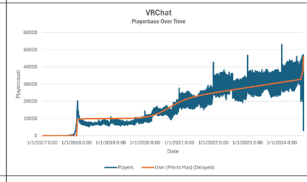
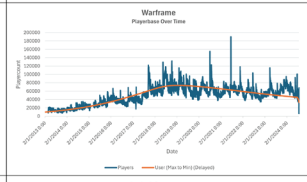
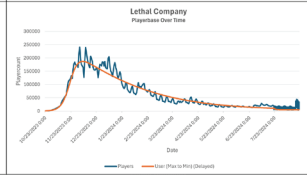
Game	Picture	τ (Short Term)	τ (Long Term)	Δ (Short Term)	Δ (Long Term)
Call of Duty (2022) *Spike and Decline Profit Driven*		20	N/A	25	N/A
Destiny 2 *Spike and Decline Profit Driven*		50	N/A	1	N/A
Lost Ark *Spike and Decline Profit Driven*		125	N/A	5	N/A
Player Unknown's Battlegrounds *Spike and Decline Profit Driven*		250	N/A	250 25 (delay increase)	N/A
Overwatch 2 *Spike and Increase Profit Driven*		10	500	1	N/A
Hunt: Showdown *Sustained Growth Profit Driven*		1000	3000	750	2000
War Robots *Sustained Growth Profit Driven*		N/A	900	500	N/A
DOTA 2 *Growth and Decline Profit Driven*		N/A	900	200	950
Apex Legends *Growth and Decline Profit Driven*		500	2000	60	N/A
Tom Clancy's Rainbow Six Siege *Growth and Decline Profit Driven*		N/A	1000	950 170 (delay increase)	N/A

Table 3: Real and predicted graph data for each community-driven game, along with the graph behavior and values produced from the equations.

Game	Picture	τ (Short Term)	τ (Long Term)	Δ (Short Term)	Δ (Long Term)
Counter Strike 2 *Sustained Growth Community Driven*		50	4200	1200	3000
Dead by Daylight *Sustained Growth Community Driven*		2000	3700	100	2000
Rust *Sustained Growth Community Driven*		3500	200000	1	1
Team Fortress 2 *Sustained Growth Community Driven*		70	9000	1300	4500
War Thunder *Sustained Growth Community Driven*		N/A	1100	2300	N/A
VRChat *Sustained Growth Community Driven*		25	1200	370	1300
Warframe *Growth and Decline Community Driven*		N/A	1200	750	2000
Lethal Company *Spike and Decline Community Driven*		50	N/A	30 5 (delay increase)	N/A

Discussion

The data has two added labels: "Community Driven" and "Profit Driven". The way that "Community Driven" was reached was through research to determine if these games supported community-created content or had a strong community base that was actively involved in the development of the game. Otherwise, it was labeled "Profit Driven" as the community would only have minor involvement in the development and content creation of the game, and/or the game's

DOI: 10.36838/IJHSR88.76

progression was affected by inherently predatory microtransactions.

Games like Apex Legends, Overwatch 2, and War Robots were all considered "profit-driven" as the games did not allow for direct community content contributions and/or had many rather predatory microtransactions that dampened user experience and/or provided an unfair advantage.

Games such as Lethal Company, VRChat, and Team Fortress 2 were all considered "Community Driven" as the games allowed and endorsed extensive community content creation and modding tools to help improve and enhance the experience of the game for everyone.

A common pattern that can be seen is that profit-driven games tend to decrease in player base over time, while community-driven games increase in player base over time. There are a few outliers, such as War Robots, Overwatch 2, and Lethal Company. A reason for both profit-driven games seemingly increasing in popularity is that the population data was taken from only one platform, Steam. The main user base of War Robots is located on mobile platforms, such as iOS and Android, and the main user base of Overwatch 2 is located on Battle.net, Blizzard Entertainment's game store and launcher. Therefore, the increasing user base on Steam can be attributed to a platform conversion, where the user base on one platform is slowly being converted to another. Lethal Company, despite being community-driven, still declines in popularity, which can be explained by its simple premise, leading to fatigue and a lack of progression for the player. Even after its brief period of extreme popularity, the game still possesses a solid core fanbase and modding community.

Game theory can be used strategically to explain the other parts of this data. Figure 2 shows the three different types of game theory games that have been covered, how they align when put up with each other, and what sort of categories they fall into.

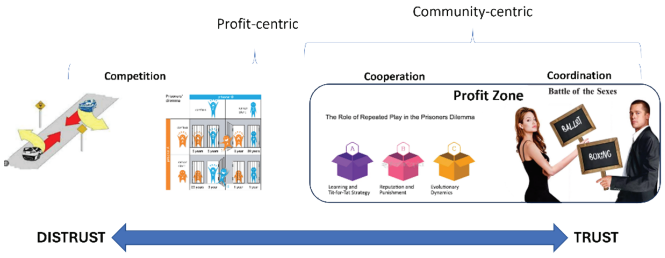


Figure 2: A graphical depiction of the game theory graphs and their relation to community-based and profit-based business strategies.

The Game of Chicken and Prisoner's Dilemma games on the left are grouped with an arrow pointing to the left and represent profit-oriented games. There is more competition and potential for conflict between the developer and the consumer, meaning that player interest and player numbers may drop, though the initial engagement may be high.

The Prisoner's Dilemma and the Battle of the Sexes games on the right are grouped with an arrow pointing to the right and represent community-oriented games. There is more cooperation and potential for connections to be formed between the developer and the consumer, meaning that player interest and player numbers may be steadier and increase over time, though the initial engagement may start low.

It may be noted that the Prisoner's Dilemma is in both the community-oriented and profit-oriented categories. This is because the Prisoner's Dilemma is an interesting case where only playing a game once will most often lead to competition since the best option is to compete due to lack of trust; however, playing a game multiple times will often lead to cooperation, since the much better option, in the long run, is to cooperate.

While most businesses align with an extreme end of the spectrum, it is possible to strike a balance in the middle, where profit can be earned and is stable, but the community can also be formed and endorsed. It is difficult to find this balance, but with the correct adjustments to both a business and a game, developers can keep players engaged in their game and keep profits stable as the game stays in service.

The business models featured in "The Profit Zone" also support this ideology, as they all encourage the creation of connections between developers and/or players – such as the switchboard and setting-the-standard models – or encouraging loyalty from the player through accessibility and ease of use – such as the product pyramid and total solution models – forming communities that last longer and are healthier.

Conclusion

In summary, game theory can be used to logically improve modern-day strategies in the video game industry. By analyzing player base data from different games, we drew the conclusion that a healthier game for both players and developers puts player sentiment into account to increase the game's longevity and hopefully make the game as community-oriented as possible, building up connections and attachments between the players and developers.

This can be shown by how community-oriented games usually experience a slower yet more stable increase in users, whereas profit-oriented games experience large spikes and plateaus because community-oriented games strive for longevity and community health, while profit-oriented games tend to prioritize monetization.

Game theory can be used to come to logical conclusions as to why each decision leads to each outcome, since prioritizing profit only works effectively in the short term, while prioritizing community leads to much more substantial long-term profits, as shown in the Prisoner's Dilemma, where "cheating" is short-term and repeated "cooperating" is long-term. This, therefore, leads to an average net positive in profits when a game is community-oriented, as players stay longer and spend more over that time.

A business can become more community-oriented by listening to player feedback for all parts of their business model, as well as engaging their community in the evolution of the product. Allowing developers and players to work together to improve a game for all involved should be increasingly normalized in today's video games industry, as more and more innovative and unique games are created and released, yet undermined by poor management.

Ultimately, due to the prevalence of subscription models, the results of this analysis can be utilized in more than just the video game industry. The logic and reasoning done through game theory can be used to explain any decision made in almost every other business.

References

1. Slywotzky, Adrian J. The Profit Zone: How Strategic Business Design Will Lead You to Tomorrow's Profits. Crown/Archetype, 2007.
2. Electronic Arts: Business Model, SWOT Analysis, and Competitors 2024 - PitchGrade. https://pitchgrade.com/companies/electronic-arts. Accessed 19 Apr. 2025.
   Pereira, Daniel. Steam Business Model. 23 June 2022, https://businessmodelanalyst.com/steam-business-model/.
3. Cuofano, Gennaro. "Inside The Epic Games Empire And Business Model." FourWeekMBA, 20 June 2024, https://fourweekmba.com/epic-games-business-model/.
4. Cuofano, Gennaro. "The Nintendo Business Model In A Nutshell." FourWeekMBA, 5 Apr. 2024, https://fourweekmba.com/nintendo-business-model/.
5. Sothmann, Anton. Business Models in the Video Game Industry. 2018. Aalto U, MS thesis. Aaltodoc, Aalto University, aaltodoc.aalto.fi/server/api/core/bitstreams/fede4ca4-d1b1-4f6e-ace2-1393f4c3b81d/content. Accessed 6 Mar. 2024.
6. Papayoanou, Paul A. Game Theory for Business: A Primer in Strategic Gaming. Gainesville, Probabilistic Publ., 2012.
7. Friman, Eric Fredriksson, and Ola Zätterlund. The Dark Patterns of Battle Passes: Investigating player attitudes to a growing type of microtransaction. 2023. UMEÅ U, MA thesis. Simple Search, Uppsala University, umu.diva-portal.org/smash/get/diva2:1776593/FULLTEXT01.pdf. Accessed 6 Mar. 2024.
8. King, Daniel L., Delfabbro, Paul H., Gainsbury, Sally M., Dreier, Michael, Greer, Nancy, and Billieux, Joel. "Unfair play? Video games as exploitative monetized services: An examination of game patents from a consumer protection perspective." Computers in Human Behavior, vol. 101, Dec. 2019, pp. 131-43. ScienceDirect, https://doi.org/10.1016/j.chb.2019.07.017. Accessed 6 Mar. 2024.
9. Alha, Kati, Kinnunen, Jani, Koskinen, Elina, and Paavilainen, Janne. "Free-to-Play Games: Paying Players' Perspective." Mindtrek '18: Proceedings of the 22nd International Academic Mindtrek Conference, pp. 49-58, https://doi.org/10.1145/3275116.3275133. Accessed 15 Dec. 2023.
10. Anderson, Eugene W., Fornell, Claes, and Rust, Roland T. "Customer Satisfaction, Productivity, and Profitability: Differences between Goods and Services." Marketing Science, vol. 16, no. 2, May 1997, pp. 129-45, https://doi.org/10.1287/mksc.16.2.129. Accessed 15 Dec. 2023.
11. Chen, Yu, Duan, Haihan, and Cai, Wei. "The Advertising in Free-to-play Games: A Game Theory Analysis." GameSys '21: Proceedings of the Workshop on Game Systems (GameSys '21), https://doi.org/10.1145/3458335.3460812. Accessed 15 Dec. 2023.
12. Goh, Edward, Al-Tabbaa, Omar, and Khan, Zaheer. "Unravelling the Complexity of the Video Game Industry: An Integrative Framework and Future Research Directions." Telematics and Informatics Reports, vol. 12, Dec. 2023, p. 100100. DOI.org (Crossref), https://doi.org/10.1016/j.teler.2023.100100.

Author

Richard L. Wan is an 11th grader at Central Bucks High School East. This paper was inspired by a science fair project in which he used game theory to prove that specific business strategies led to more satisfaction amongst customers while not impacting profits negatively.

DOI: 10.36838/IJHSR88.76

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

CFD Aerodynamic Analysis of Design and Force Coefficients of 2022 and 2026 Formula One Cars

Yuto Sakamoto

ACS International, 61 Jln Hitam Manis, Singapore, 278475, Singapore; yuto.sa0321@gmail.com

ABSTRACT: The 2026 Formula One car is different from any other car that Formula One (F1) has introduced in previous eras; it is, in fact, the entire car that is visually unique. In this investigation, the aerodynamics of the 2026 and 2022 F1 cars are compared to analyze aerodynamics by using a CFD, incompressible flow simulation provided by SimScale, with 3D car models available online (both are unofficial models). This research aims to compare aerodynamic characteristics of the 2022 and 2026 Formula One regulation cars and examine possible effects on racing performance. Results show that the 2026 model has eliminated the ground effect, indicating a major loss of a source of downforce. Simultaneously, the upward diversion of wake behind the rear wing is reduced, potentially compromising overtaking opportunities for the car behind. CFD analysis concludes that the 2026 regulation reduces both the lift coefficient and the drag coefficient by approximately 75% compared to the 2022 model. These findings demonstrate the substantial reduction in downforce and altered aerodynamic characteristics of the 2026 Formula One car, indicating that the car compromises cornering performance and race competitiveness.

KEYWORDS: Engineering Mechanics, Aerospace and Aeronautical Engineering, Computational Mechanics, Racing Car, F1.

Introduction

Formula One began its journey on May 13, 1950, at the Silverstone circuit in England.[1] Ever since the first grand prix, Formula One cars have been evolving as a pinnacle of motorsports with their engine development, energy efficiency, and safety to pursue speed. One of the remarkable and perhaps revolutionary developments of F1 is the aerodynamics. In the late 1970s, 20 years after the inception of F1, the concept of innovative aerodynamics, with inverted wings, was introduced by the Lotus 49. In 1977, another aerodynamic concept of ground effect was introduced by the Lotus 78.[2] The principal purpose of aerodynamics in F1 is to minimize drag and increase downforce of the car, achieving incredible speed not just in the straight but also in corners.

Once every few years, the FIA (Federation Internationale de l'Automobile) introduced new concept cars, which became the basis for other teams' foundations for the forthcoming seasons. Recent developments include major aerodynamic regulation changes in 2009, the introduction of new hybrid-turbo engines in 2014, and the complete resurrection of ground-effect cars in 2022. In 2026, the FIA launched an overhaul of the F1 car, featuring a futuristic design and an utterly and fundamentally new concept. The 2026 car is a completely unexplored character; even the teams have no solid certainty of what to expect from it. There is every reason to be excited about what the 2026 car holds and how it will redefine the competition in the car market.

In this research, the aerodynamic performance of the 2026 car is investigated using computational fluid dynamics (CFD), with a comparison to the 2022 model to evaluate the effects of the new regulations on airflow behavior, downforce generation, and ground effect in particular.

Background information

General Aerodynamic Performance of a Formula One Car:

The average maximum speed of a Formula One car is between 210 and 220 mph.[3] At this speed, without adequate downforce, the Formula One car would be flying everywhere. Downforce is mainly generated by the 2 wings, the front wing and the rear wing. Essentially, these wings function as inverted wings.

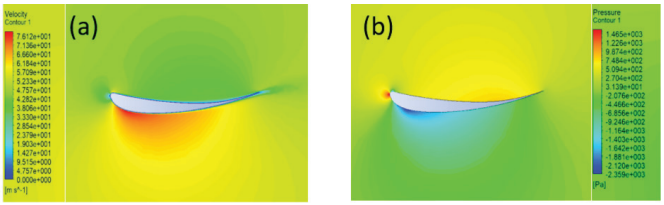


Figure 1: This shows the aerodynamics of an inverted airfoil.[4] Velocity (a) and pressure (b) contours around an inverted airfoil are shown. The red region indicates high velocity underneath the airfoil, which corresponds to lower air pressure shown in (b).

The generation of downforce can be explained by Bernoulli's principle, which states that as the velocity of airflow increases, the pressure decreases for a single, steady incompressible flow. Essentially, the airfoil would have two independent streams of air: one across the top surface and the other across the bottom surface. Furthermore, the inverted airfoil has more curvature at the bottom surface, whereas the curvature is milder across the top surface.[5] This airstream moves down along the curvature; to prevent a void, the airstream is pulled towards the trailing edge of the airfoil, accelerating the airstream upwards.[5] As a result, according to Bernoulli's principle and greater curvature, the airstream at the bottom surface has relatively higher velocity, hence, lower air pressure. This is illustrated in Figure

DOI: 10.36838/IJHSR88.81

1, where the region of higher air velocity (near the bottom surface) has lower air pressure, and vice versa. In contrast, the airstream at the top surface has relatively lower velocity, hence, higher air pressure. This results in a pressure gradient between the surfaces, creating a down force, effectively "pushing down" the airfoil. Bernoulli's equation is explained in terms of the conservation of energy. The sum of pressure energy, potential energy per unit volume, and kinetic energy per unit volume is always constant in a single streamline. This can be used to define downforce. For the top airstream:

$$P_{Top} + pgh_{Top} + \frac{1}{2}\rho\left(V_{Top}\right)^2 = P_{\infty} + \rho gh_{\infty} + \frac{1}{2}\rho\left(V_{\infty}\right)^2$$

where the subscript ∞ represents the free stream condition. However, because both the freestream and the airstreams around the airfoil are almost at the same altitude, this allows it to cancel both "ρgh"s. Therefore, the pressure difference between the airstreams is

$$P_{Top} - P_{Bottom} = \frac{1}{2}\rho\left(V_{Top}\right)^2 - \frac{1}{2}\rho\left(V_{Bottom}\right)^2$$

$$\Delta P = \frac{1}{2}\rho\left(V^2{}_{Top} - V^2{}_{Bower}\right)$$

Since this pressure difference generates the downforce,

$$L = \frac{1}{2}\rho A\left(V^2{}_{Top} - V^2{}_{Bottom}\right)$$

In addition, downforce is dependent on the lift coefficient, a dimensionless constant that indicates the airfoil's capability to generate lift. However, in this case, the constant has a negative value, indicating that it is downforce rather than lift.

$$L = C_L \cdot \frac{1}{2}\rho A\left(V^2{}_{Top} - V^2{}_{Bottom}\right)$$

$$C_L = \frac{2L}{\rho V^2 A}$$

where:
L: downforce in N,
ρ: density in kgm^{-3}
V: velocity of freestream in ms^{-1}
A: the wing area in m^2.

DRS

However, the rear wing of Formula One is not just a rear wing; it can also deploy a DRS (drag reduction system). The DRS was introduced to F1 cars in 2011 to promote wheel-to-wheel racing by decreasing downforce and drag force, thereby enhancing the car's speed. DRS contributes to a 53% reduction in the lift coefficient of the car, as well as a 78% decrease in the drag coefficient of the car.[8]

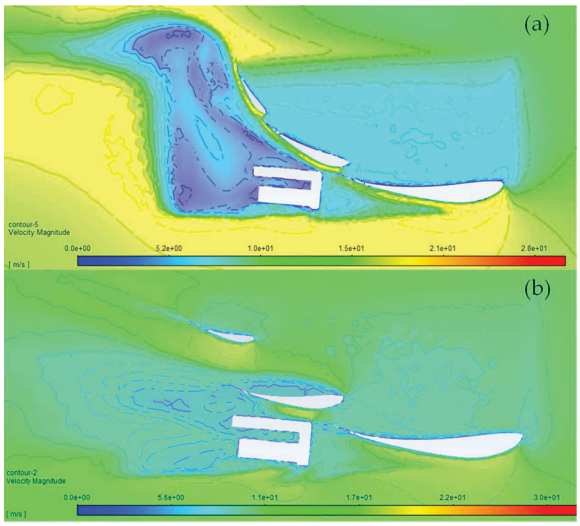


Figure 2: Aerodynamic effect of DRS.[8] Velocity contours of airflow around the rear wing with DRS closed (a) and open (b) are shown. In (a), the large blue region indicates low velocity behind, and when DRS is opened, high velocity is shown just behind the airfoil.

Upon activation of DRS, the multi-element rear wing opens until it reaches the angle of attack of 0 degrees.[8] The Figure 2(a) consists of a large region of low velocity (meaning high air pressure) behind the rear wing. This ultimately contributes to pressure gradients that generate downforce. In contrast, by allowing the airstream to pass between the airfoil, the air velocity in Figure 2(b) has not changed significantly, and the low and high air pressure regions that promote downforce have disappeared; lower downforce results in lower induced drag. Moreover, since the surface area that was getting "pushed" by the receiving air velocity is reduced, the drag force is reduced. Lower drag and downforce allow the car to maximize its pure speed. DRS is mostly applied on long straight lines in the race circuit.[9]

2022 Car:

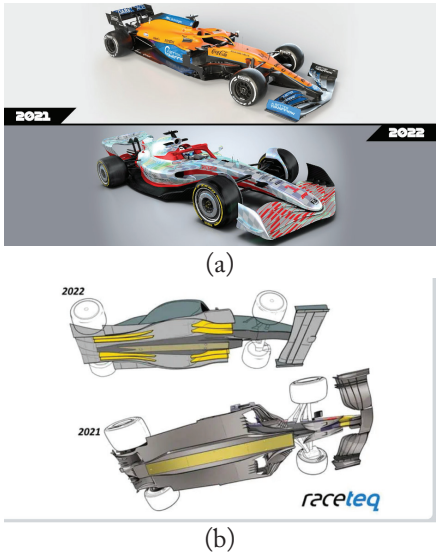


Figure 3: Car floor design difference between the 2022 car and the 2021 car (Giuliana).[11] The images show a comparison between the side and the underside of the car. The 2021 car has an angular body with a simpler chassis. On the other hand, the 2022 car has a smoother body and introduced venturi tunnels in the chassis.

Ground Effect:

The 2022 car was revolutionary, with a futuristic design compared to the previous generation's F1 cars. The car was specifically designed to facilitate further overtakes in racing and is a revival of the ground effect car.[12] Ground effect is an aerodynamic phenomenon that occurs near the ground, resulting in the creation of a low-pressure zone underneath the car, which generates downforce. The current ground effect from F1 cars arises from the Venturi effect, stating that the velocity increases in a constricted area (hence, lower pressure).[13] This ground effect is mostly provided by the chassis of the car, more specifically, the venturi tunnels.[13] This means that the 2022 cars can carry high speed and traction in the corners. Although the incredibly low pressure underneath the car sucks the car down extremely close to the ground, that venturi tunnel becomes a vacuum. This triggers the sudden disappearance of the downforce, as a result, the car springs up only for the venturi tunnel to reopen.[14] This repeated oscillation-like movement due to extreme ground effect is called porpoising and was a major issue to be solved at the beginning of the 2022 season.[14]

Promoting Better Racing:

Another characteristic of the 2022 car is that it has rounded edges on the wings. Previous designs diverted the flow upwards and outwards, meaning the car behind received all the turbulence (or "dirty air"), deteriorating drivability.[15] This new design permits narrower vortices, rotational flow of air to be created, collecting turbulence and wake, and diverting them upwards.[16] Thus, a driver behind experiences less dirty air and overtaking becomes less demanding.

2026:

Figure 4: The 2026 car was announced via FIA.[17] The new regulation introduces new features to the cars, such as X-mode and Z-mode.

The current 2022 generation of F1 is coming to an end in 2025; instead, the 2026 generation will begin. As observed from the preliminary design of the 2026 car, its appearance exhibits a significant difference from that of the 2022 car.

Size:

The 2026 car will be scaled down, with reduced wheel size and a lower car floor. Cutting the wheel size significantly reduces overall weight, contributing to a total of 30 kg in car weight reduction.[17] The weight of the car is important in racing; a lighter mass allows the car to accelerate quickly, resulting in higher energy efficiency to generate the force required to move the car. Furthermore, the light weight decreases the moment of inertia, meaning turning the car becomes easier. Improving the agility of the car is one of the primary objectives of the 2026 car, aiming to enhance the car's responsiveness for competitive racing.

Z-mode & X-mode:

The 2026 car features a system similar to DRS, along with new features called Z-mode and X-mode, a 100 *mm* narrower two-element front wing, and a three-element rear wing with the lower beam removed.[17]

Z-mode: This is the standard mode in which elements of the front and rear wings open at an angle to aid cornering speed.[17]

X-mode: Deploys a low drag configuration, opens both front and rear elements to maximize straight line speed.[17] This is an utterly new feature, as only the rear wing element was subject to DRS.

■ Methodology

Using computational fluid dynamics (CFD), the aerodynamics of the 2022 and 2026 model cars are analyzed. The 2022 car, publicly published by ss10suryansh, and the 2026 car, by reesewu2005, will be utilized for this research.

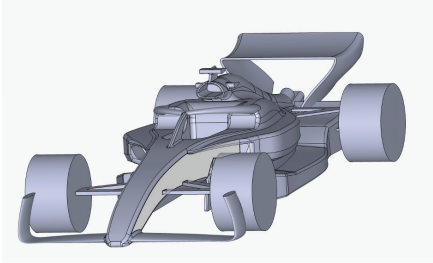


Figure 5: 3D Design of the 2022 model.

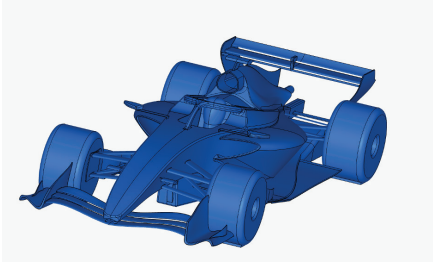


Figure 6: 3D model of the 2026 model. Figures 5 and 6 are the 3D models of the 2022 and 2026 cars used for simulation. The 2026 model also includes the driver's head.

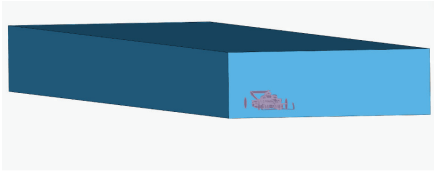


Figure 7: Flow region geometer (model car inside, in red). The flow boundary is defined around the car with the model at the center, where the wind inlet is set on the surface just in front of the model, and force sensors are on the surface of the model.

Before the CFD simulation, it is essential to establish the correct setup conditions for accurate and reliable results. The setup process involves defining the flow region, specifying boundary conditions, and setting appropriate parameters to replicate real-world scenarios. There is a slight difference in

DOI: 10.36838/IJHSR88.81

the setup of the two cars: for the 2026 car, the car model and airflow region are independent of each other, while for the 2022 car, the car model is embedded within the airflow region itself. A mesh is a network of cells and nodes that allows analysis of a simulation with complex geometry. The 2026 model consists of 25.8 million cells and 7.8 million nodes, with a gap refinement factor of 0.05 and a runtime of 1.8e+4 seconds. The 2022 model consists of 3.1 million cells and 3.8 million nodes in the internal meshing model, which is run by the Hex dominant algorithm, with a runtime of 1.8e+4 seconds.[18] The observed discrepancy in mesh structure arises from the original author employing different meshing parameters across the models.

A. Creating an enclosure that represents the flow region. The flow region simulates a controlled region, akin to a wind tunnel, where the interaction between airflow and an object inside can be analyzed. Within this enclosure, various boundary conditions and surface sensors are assigned.

B. The wall in front of the model car is where the velocity inlet is assigned. This is also where air is introduced into the flow region, with a density of 1.196 $kgms^{-3}$, corresponding to the air density at 20 °C. The airflow has a velocity of 50 ms^{-1}.

C. The opposite wall to the velocity inlet is defined as the pressure outlet. This boundary condition permits the airflow to exit the domain, maintaining a realistic pressure gradient within the domain. The ground wall, which simulates a road or racetrack, is set to move at a velocity of 100 ms^{-1} along the x-axis. Furthermore, wheel rotation is an essential factor in analyzing aerodynamic effects around the tires. Both front and rear wheels have a rotational velocity of 166 radians per second.

Results

Qualitative Analysis – 2022:

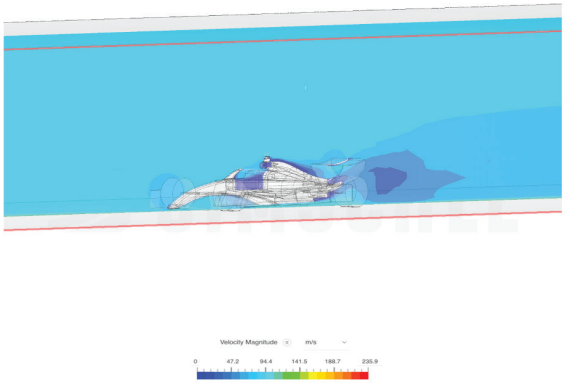


Figure 8: The velocity magnitude of the air around the 2022 model-1. The results show a smooth, continuous, and upcurved lower velocity region behind the 2022 car. Slightly darker blue indicates a slower velocity compared to the light-blue free stream of air.

Further evidence that the 2022 F1 car's effectiveness in preventing dirty air is shown in the velocity magnitude of the free stream air. As shown in the scale below, the closer the color is to dark blue, the slower the velocity of the free stream. The region near the halo or cockpit of the car has the darkest area as air particles collide with the wall, thereby slowing down. The area directly behind the car also has a region of slower air particles, as less air is diverted into the area. The importance of this solution field compared with the vorticity solution field is that the air stream is diverted upward after the particles pass through the car. This is one of the unique characteristics of the 2022 F1 car, which directs the air upwards immediately after the rear wing to create a region that prevents dirty air.

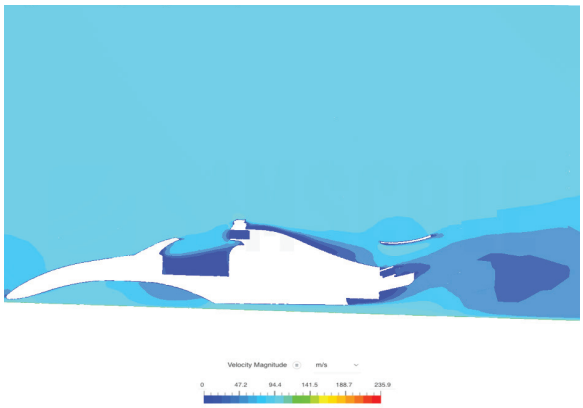


Figure 9: Velocity magnitude around the 2022 model-2. The region between the chassis of the car and the floor has a sky-blue color, indicating that the velocity of the air stream is as high as that of free stream air.

From the earlier Bernoulli's equation states that

$$P_{top} - P_{bottom} = \frac{1}{2}\rho\left(V_{top}\right)^2 - \frac{1}{2}\rho(V_{bottom})^2$$

This equation indicates that an increase in velocity results in a decrease in pressure. Applying Bernoulli's principle in this context, the higher velocity near the ground causes lower pressure in this region compared to the region above the F1 car. This pressure gradient between the top surface of the car and the bottom surface of the car creates a net downward force. In other words, this represents the ground effect of the 2022 F1 car.

Quantitative Analysis – 2022:

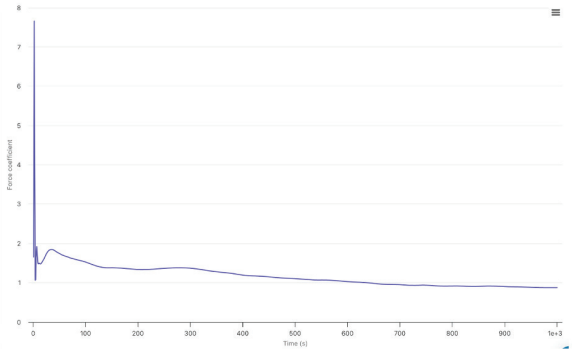


Figure 10: Graph of the drag coefficient of the 2022 model. This graph shows the drag coefficient rapidly decreasing from an initial fluctuation and then gradually stabilizing to a value closer to 1.0.

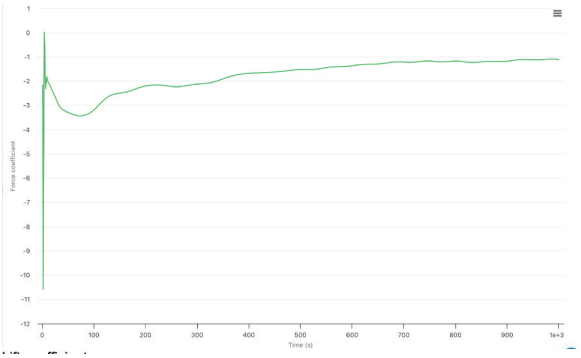


Figure 11: Graph of the lift coefficient of the 2022 model. The value of the lift coefficient gradually stabilized around -1.0. This indicates that the 2022 car constantly generates downforce once the airflow reaches equilibrium.

Both Figure 10 and Figure 11 are collected from the simulation run of the 2022 model on SimScale. The x-axis shows the times, whereas the y-axis shows the force coefficient, presenting the car's coefficient at each second after the simulation began. There are fluctuations in the data at the early part of the graph, but these are mostly due to high unsettled air passing through the car, which gives momentarily high or low coefficients and is not reliable. However, towards the end, there is a tendency for the values to stabilize (in other words, the gradient of the graph approaches zero). This is a more reliable value of the coefficient, as this indicates that the airflow around the car has reached a relatively stable state. Hence, the last value observed (terminal value) on the graph will be used to calculate lift and drag forces.

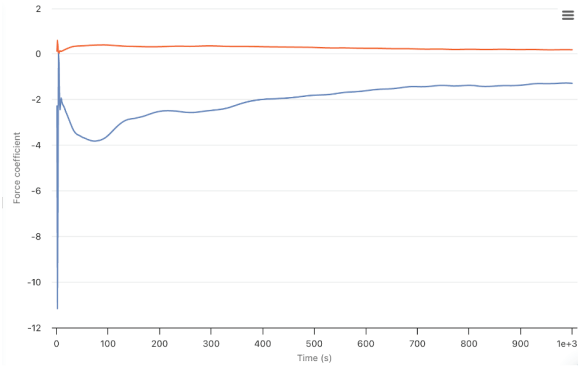


Figure 12: The graph shows the lift coefficient of the front of the model, C(f), and the rear end of the model, C(r). Both values of Cl(f) and Cl(r) seemed to settle around a constant value.

Table 1: This Table shows the terminal values of Cl(f) and Cl(r) of the 2022 model obtained from Figure 12. Positive value indicates that it is contributing to lift, while a negative value indicates it is contributing to downforce.

Cl(f)	Cl(r)
0.1872	-1.2864

Cl(f) and Cl(r) are the measures of the lift coefficient generated from the front or rear section of the car, respectively. From Table 1, the positive value of the lift coefficient for the front area indicates that this area contributes to generating lift rather than downforce. This suggests that the aerodynamic profile of the front end contributes minimally or even detrimentally to vertical load generation, potentially damaging the front tire grip at corner entries. On the other hand, Cl(r) has a significantly larger negative value. Hence, this suggests that the downforce of the 2022 model is mostly provided by the rear end of the car.

Table 2: This table presents the terminal values of the Drag coefficient and Lift coefficient of the 2022 model obtained from Figure 10 and Figure 11, respectively.

Force coefficient	Terminal value
Drag coefficient	0.8749
Lift coefficient	-1.099

Using the force coefficients values, both the Drag and Lift forces can be calculated using the formulas:

$$C_L = \frac{2L}{\rho V^2 A}$$

$$C_D = \frac{2F_D}{\rho V^2 A}$$

Where:
L: force
F_D: Drag force
C_L: Lift coefficient
C_D: Drag coefficient
ρ: Air density (1.195 $kgms^{-3}$)
V: Velocity of free stream air (50 ms^{-1})
A: Surface area

Table 3: This table shows forces on the 2022 model calculated using the force coefficients listed in Table 2.

Forces	Terminal value (N)
Drag force	1306.88 x area
Lift force	-16434.005 x area

Qualitative Analysis – 2026:

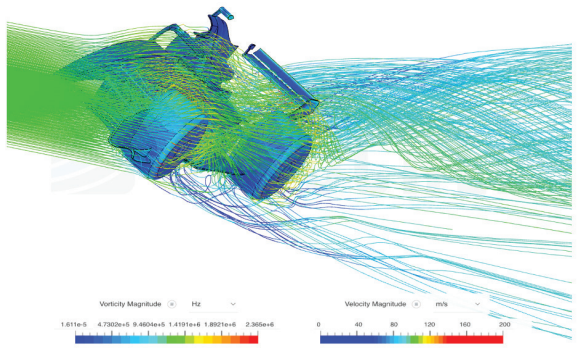


Figure 13: Diagonal view particle trace visualization of the 2026 model. Trace around the vehicle shows an entrance velocity of 100 ms^{-1}, and vorticity is depicted on the surface of the car.

DOI: 10.36838/IJHSR88.81

This solution field represents the magnitude of vorticity of Formula 2026 in a free-stream car at 50 ms^{-1}. The coloring of velocity is on the air particles, and the coloring of vorticity magnitude is on the F1 car. Similar to that of a 2022 car, the velocity of the air stream decreases when a free stream hits the car. However, the difference between the 2022 car is that the intense undulation is observed around the front tires and immediately after the rear wing. The vorticity is higher at the edges of the car component, most obviously shown on the rear wing.

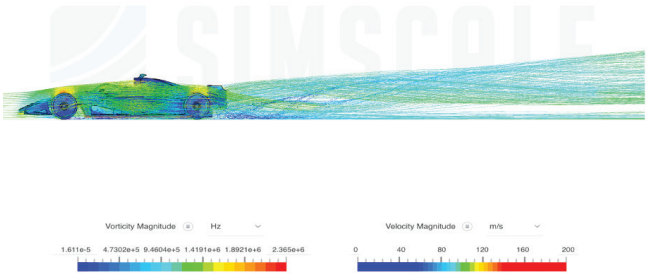


Figure 14: The figure shows a side-view particle trace visualization of the 2026 model. Particle trace shows that the air flows smoothly and gradually above.

From the horizontal view (Figure 14), the effect of dirty air on the car behind is clearer. Unlike the 2022 car, where the air is diverted upwards immediately, the undulating air remains at the ground level. This indicates that the car behind experiences the effect of vortices.

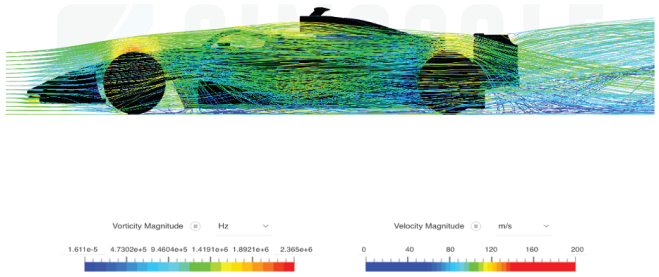


Figure 15: The diagram shows the particle trace around the 2026 model (focused). While higher velocity regions are concentrated above and behind the car, the lower velocity region is observed beneath the car.

The color on the particle represents the velocity magnitude, in which, as the color approaches red, it indicates that the velocity is higher. The air particles below the chassis are blue, indicating a low velocity of air, which means it is a region of high pressure. On the other hand, the color of particles on the top surface is mostly green, suggesting a higher velocity of air and a low-pressure region. Hence, there is a higher-pressure region beneath the car and a lower-pressure region above it. The pressure gradient is set such that there is a net force acting upwards. This indicates that ground effect is undetectable for the 2026 model, resulting in the loss of downforce typically gifted by ground effect in 2022.

Quantitative Analysis – 2026:

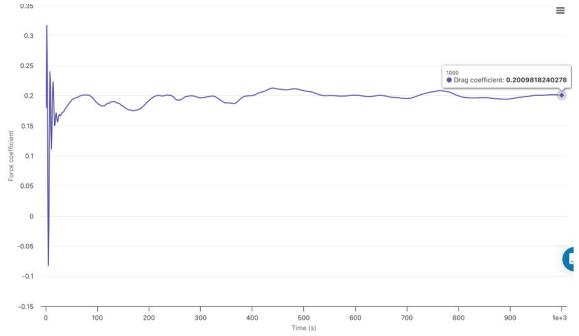


Figure 16: The graph shows the drag coefficient of the 2026 model. The value mostly fluctuates around 0.2.

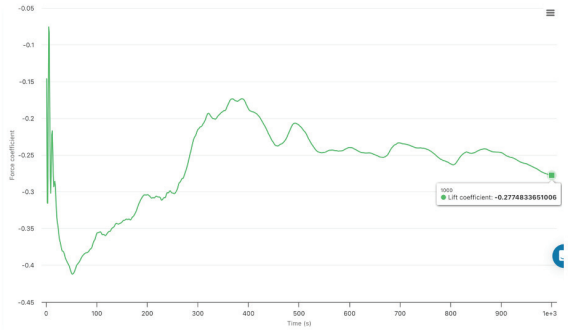


Figure 17: The graph shows the lift coefficient of the 2026 model. The dataset demonstrates the greatest variation in the magnitude of force coefficients, which are noticeably larger than those observed in other graphs.

Figures 16 and 17 show the variation of the lift coefficient over time for the 2026 model. In contrast to the 2022 model, frequent fluctuations in data are observed. Overall, the drag coefficient ranges between 0.15 and 0.2, which is significantly lower than the 2022 model, which ranges from 0.8 to 2. The lift coefficient of 2026 exhibits even more frequent fluctuations in value, ranging from -0.17 to -0.278, which is significantly lower in magnitude compared to the lift coefficient of the 2022 car. In conclusion, both drag coefficient and lift coefficient have a lower magnitude of 2026, which is significantly lower than in 2022.

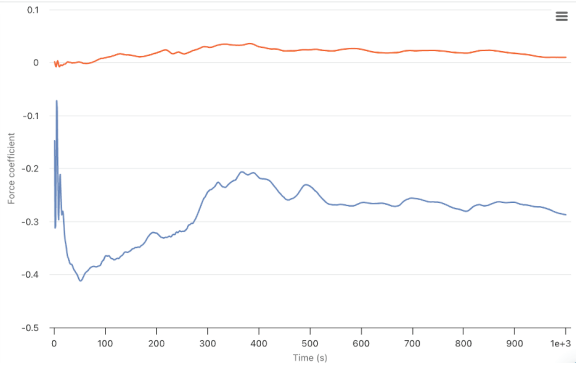


Figure 18: The graph shows Cl(f) (orange) and Cl(r) (blue) of the 2026 model. While the trend line of Cl(f) is relatively stable, the trend line of Cl(r) shows variation, which was not seen in the graph for the 2022 model.

Table 4: Table showing terminal values of Cl(f) and Cl(r) of the 2026 model from Figure 18.

Cl(f)	Cl(r)
0.0097	-0.2871

According to Table 4, the 2026 model is an aerodynamically rear-end-dominant car in terms of downforce generation. However, in comparison to the 2022 model, the front wing contributes less lift, enhancing downforce on the front section of the car. Nonetheless, downforce generated by the rear end is significantly reduced. Overall, the 2026 car is a more rear-biased vehicle, as shown in Table 4. This imbalance in downforce makes the 2026 model an aerodynamically rear-end dominant car for downforce. However, in comparison to the 2022 car, the front wing contributes less lift, thereby enhancing downforce on the front section of the car. Nonetheless, the downforce generated by the rear end is significantly less. Overall, the 2026 car is more rear-biased, and such an imbalance of downforce may cause extraordinary instability in corners.

Table 5: Table showing terminal values of the Drag coefficient and Lift coefficient of the 2026 model from Figures 16 and 17, respectively. Compared to Table 2, the magnitude of force coefficients is significantly lower for the 2026 model.

Force coefficient	Terminal value
Drag coefficient	0.20098
Lift coefficient	-0.27748

Applying the same formula to calculate the two forces (Table 6),

Table 6: The table shows forces on the 2026 model calculated using the force coefficients listed in Table 5.

Forces	Terminal value (N)
Drag force	300.4561 x area
Lift force	-414.8326 x area

Comparison between the forces of the 2022 and 2026 F1 cars

Table 7: The table presents a comparison of forces on the model between 2022 and 2026. Overall, without accounting for area in the calculation, the 2022 model experiences greater Drag force and Lift force compared to the 2026 model.

Forces / Model of F1 car	2022 (N)	2026 (N)	Reduction
Drag force	1306.63 x area	317.9043 x area	75.66%
Lift force	-1641.6312 x area	-414.8326 x area	74.7%

From Table 7, it is conclusive that 2026 has both a lower drag force and downforce. This reduction is indeed significant, with a 75.66% decrease in drag force and a 74.7% decrease in downforce for the lift force for the 2026 model compared to the 2022 model. This is slightly contradictory to the statistics from the FIA, which claimed a 50% reduction in drag force and only a 30% reduction in downforce.[4] The difference in the statistical analysis is partially due to the unknown surface area. However, because the entire size of the 2026 car is smaller than that of the 2022 car, it is less likely that the surface area of the 2026 car is greater than that of the 2022 car. In other words, the reduction of force compared to 2022 is less likely to decrease (e.g., it cannot be lower than 75.66% since the ratio of surface area will never exceed 1). Although the 2026 model shows a reduction in downforce in percentage terms, the lift coefficients of both models are closer to those of conventional road cars, suggesting potential limitations in the modelling approach or simulation conditions. Nonetheless, it is also unclear how FIA determined that there is a reduction or in what context this reduction occurs.

Discussion

Overall, the data analysis using SimScale confirmed that the 2026 car generates significantly less downforce than the 2022 model. This reduction is evident both qualitatively and quantitatively. In particular, the distinctive aerodynamic features that contributed to the performance and stability of the 2022 car—such as the utilization of ground effect and the upward diversion of airflow behind the rear wing—are no longer present in the 2026 design.

The 2022 model takes full advantage of ground effect, as shown by the high-velocity airflow between the car's floor and the ground, resulting in low pressure beneath the chassis. According to Bernoulli's principle, this creates a pressure differential that generates significant downforce, keeping the car pressed to the track and enhancing grip during high-speed cornering. This improves lap times and drivability of the car while maintaining higher cornering speeds.

Moreover, the airflow is diverted upward after the rear wing, reducing the intensity of the dirty air experienced by a car following behind, which enables the trailing car to accelerate more quickly. This is a key design principle introduced in 2022 to promote closer racing and overtaking during the race.

In contrast, the 2026 car lacks these aerodynamic mechanisms. As seen in the particle trace, the slow-flowing air beneath the car results in high pressure under the chassis and a net upward force, significantly reducing ground effect. Furthermore, particle traces in the vorticity field reveal significant airflow undulation and turbulent vortices, particularly behind the rear wing and around the front tires. In other words, the front tires are more prone to degradation from air resistance, suggesting that long run might not be an advantage for 2026 cars compared to other generations.

These regions of unsteady flow imply a resurgence of dirty air, which could hinder the ability of cars to follow closely. The car behind could experience a sudden loss or gain of downforce, losing control of the car, especially when braking. Importantly, the car experiences more aerodynamic drag due to this dirty air, which could slow it down in the straights. Nevertheless, the newly introduced X-mode, which can open both the front and the rear wings, may balance out the negative effects caused by dirty air by reducing greater downforce in exchange for boosting straight-line speed. From a race-craft perspective, this suggests that while following performance may marginally im-

DOI: 10.36838/IJHSR88.81

prove on straights with activation of X-mode, the loss of overall downforce could still limit close-corner following, meaning dirty air remains a significant constraint. Overall, there may be fewer overtaking opportunities for 2026 cars, contrary to what the 2022 regulations intended to provide.

In general, the 2026 model has lower drag. As a result, it may benefit straight-line speed and fuel efficiency (since less power is required to overcome aerodynamic drag), but the drastic reduction in downforce compromises performance in terms of cornering speed and braking. Plausibly, the 2026 cars may exhibit understeer behavior in corners, where the impact can be critically pronounced, particularly in higher speed corners like those at Spa (a racing circuit in Belgium). Understeer is a phenomenon in which the car turns less than intended due to loss of grip from the front tires. Since the downforce of the 2026 model is mostly contributed by the rear, a lack of grip from the front tires may cause understeer.

This trade-off could have major implications for driver confidence and racing dynamics. It also raises questions about whether the intended goals of the 2026 regulations—to improve sustainability and maintain racing quality—have been achieved. However, it should be noted that the simulation results, especially the reported percentage reductions in downforce and drag, should be interpreted with caution, as the effective surface area is not accounted for in the calculations and the models likely differ from those used by the FIA.

Conclusion

In this investigation, computational fluid dynamics were utilized to analyze the aerodynamic performance of the 2022 F1 model in comparison to the forthcoming 2026 model. This investigation revealed that, qualitatively, the 2026 model has eliminated the ground effect, indicating a major loss of a source of downforce. Simultaneously, the upward diversion of wake behind the rear wing is curtailed, potentially compromising overtaking opportunities for the car behind.

This is supported by Quantitative analysis, which calculates the results from the values of the Lift coefficient and the Drag coefficient obtained from CFD analysis. The analysis concludes that the new regulation would deliver approximately a 75% reduction in both drag and downforce compared to the 2022 and 2026 models.

Yet several key questions remain open. Will the dramatic downforce reduction truly advance the FIA's goals of sustainability without compromising cornering performance and close-quarters racing? Although a reduced drag force may require less fuel to generate a car's power, the broader sustainability impact of diminished aerodynamic performance remains uncertain. How might teams adapt suspension, ride-height, and wing angle strategies to recover lost grip?" To what extent does the opening of the front wing flap introduced in X-mode reduce drag? This research can be conducted using CFD, particularly for the study of X-mode. A car model with a front wing design featuring an open element can be used to run simulations, allowing for comparison with the closed wing and drag reduction from the rear wing, as measured by the Cl(f) and Cl(r).

For further investigation, wind-tunnel experiments and on-track telemetry (real-time data received from the car on the track) should be used to validate these CFD predictions under real-world conditions, particularly to capture transient effects such as porpoising and wake interaction.

References

1. Hooper, A. *Celebrating the 25 greatest races across F1's 75 years*. Formula 1® - The Official F1® Website. https://www.formula1.com/en/latest/article/greatest-races-best-25-f1-75-years-anniversary.3JnpCmS1AzPQLmqW1HWTxB (accessed 2025-05-15).
2. Hasanovic, V. *Formula 1 Aerodynamics - Introduction - F1technical.net*. F1technical.net. https://www.f1technical.net/features/21555 (accessed 2018-04-03).
3. Clark, A. *How fast do F1 cars go?* Red Bull. https://www.redbull.com/us-en/how-fast-do-f1-cars-go (accessed 2025-06-02).
4. Zhou, Z. Design of F1 Race Car Rear Wing Airfoil: Optimizing the Lift to Drag Ratio through Numerical Simulation. The Frontiers of Society, Science and Technology 2020 Vol. 2 Issue 12: 116-122. https://doi.org/10.25236/FSST.2020.021217.
5. Anderson, D.; Eberhardt, S. *How Airplanes Fly: A Physical Description of Lift*. https://www.cadmac.co.uk/index_htm_files/how_airplanes_fly.pdf (accessed 2025-05-15).
6. Shah, G.; Singhal, A.; Apte, R.; Dupetawalla, R. An Assessment of the Application of Bernoulli's Theorem in the Generation of Lift Force. *Journal of Student Research* 2021, *10* (3). https://doi.org/10.47611/jsrhs.v10i3.1759.
7. Benson, T. *The Lift Coefficient*. www.grc.nasa.gov. https://www.grc.nasa.gov/www/k-12/VirtualAero/BottleRocket/airplane/liftco.html (accessed 2025-03-10).
8. Loução, R.; Duarte, G. O.; Mendes, M. J. G. C. Aerodynamic Study of a Drag Reduction System and Its Actuation System for a Formula Student Competition Car. *Fluids* 2022, *7* (9), 309. https://doi.org/10.3390/fluids7090309.
9. Motorsports. *What is DRS in F1, how does it work and is it automatic?* www.motorsport.com. https://www.motorsport.com/f1/news/what-is-drs-in-f1-how-does-it-work-is-it-automatic/10437677/ (accessed 2025-06-02).
10. Reynolds, J. *ANALYSIS: Comparing the Key Differences between the 2021 and 2022 F1 Car Designs | Formula 1®*. www.formula1.com.
11. Giuliana, R. *Formula 1 aerodynamics: What are Venturi tunnels and how does porpoising affect F1 cars?* Raceteq.com. https://www.raceteq.com/articles/2024/08/venturi-tunnels-explainer (accessed 2025-05-15).
12. Stuart, G. *10 things you need to know about the all-new 2022 F1 car | Formula 1®*. www.formula1.com. https://www.formula1.com/en/latest/article/10-things-you-need-to-know-about-the-all-new-2022-f1-car.4OLg8DrXyzHzdoGrbqp6ye (accessed 2025-06-02).
13. Zhang, X.; Toet, W.; Zerihan, J. Ground Effect Aerodynamics of Race Cars. 2006. https://doi.org/10.1115/1.2110263%CD%94.
14. Zhu, Z. Cause and Analysis of 2022 Formula 1 "Porpoising." Highlights in science, engineering and technology 2023, 46, 19–27. https://doi.org/10.54097/hset.v46i.7659.
15. Stuart, G. *F1 slang explained: A beginner's guide | Formula 1®*. Formula1.com.2020 https://www.formula1.com/en/latest/article/a-beginners-guide-to-f1-slang.1Pg6tvGZ2y7u4KAnc8WXGl. (accessed 2025-06-02).
16. Guerrero, A.; Castilla, R. Aerodynamic Study of the Wake Effects on a Formula 1 Car. Energies 2020, 13, 5183. https://doi.org/10.3390/en13195183

17. FIA. *A New Era of Competition: FIA showcases future-focused Formula 1 regulations for 2026 and beyond*. Fédération Internationale de l'Automobile. https://www.fia.com/news/new-era-competition-fia-showcases-future-focused-formula-1-regulations-2026-and-beyond (accessed 2025-05-15).
18. SimScale. *Background for Hex-dominant | SnappyHexMesh | SimScale*. SimScale. https://www.simscale.com/docs/simwiki/preprocessing/what-is-a-mesh/snappyhexmesh/ (accessed 2025-06-02).

Author

Yuto Sakamoto is a senior international student at Anglo-Chinese School (International) in Singapore with a strong enthusiasm and academic interest in mechanical and aeronautical engineering. His research focuses particularly on the aerodynamics of Formula One cars, reflecting his passion for innovation and performance in motorsport.

DOI: 10.36838/IJHSR88.81

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Artificial Intelligence-Based Identification of Metastatic Targets and Natural Inhibitors in Non-Small Cell Lung Cancer

Prachet Patakula

Manthan School, Hyderabad, Telangana, 502032, India; patakula.prachet@gmail.com
Mentor: Nirupma Singh

ABSTRACT: Non-small cell lung cancer (NSCLC) is primarily lethal due to its high metastatic potential. In this study, we hypothesize to identify effective therapeutic targets and natural compounds to inhibit metastasis in NSCLC through an integrated computational framework that uses differential expression (DE) analysis, gene-interaction network analysis using mathematical modelling, virtual screening of natural compounds, and artificial intelligence. TCGA (The Cancer Genome Atlas) and GEO (Gene Expression Omnibus) were used to extract gene expression and clinical data. DE analysis was performed, and significant (p-value < 0.05) genes were identified. After validation for metastatic outcome, a final gene list of 850 genes was retained, which was analyzed using Reactome to map important pathways. STRING and Cytoscape were used for network construction and analysis. The top 10 hub genes were identified based on computed topological parameters. The genes were modelled with nonlinear ordinary differential equations (ODEs) to construct a dynamic network. The intersection of the top genes of static and dynamic networks presented 9 common genes. From the top genes, *BUB1* was shortlisted as the most important and relatively lesser-explored therapeutic target in NSCLC metastasis literature, based on a review of prior studies. Virtual Screening of natural compounds was run for *BUB1*, followed by building an artificial intelligence (AI) model for prioritization and feature-based interpretation of high-activity candidates. Thus, this study integrates gene expression data, network biology, dynamic systems modeling, and AI-augmented virtual screening to uncover regulatory hubs and therapeutic leads in NSCLC metastasis.
KEYWORDS: NSCLC, Bioinformatics, Cancer Metastasis, Natural Compounds, Protein-Protein Interactions.

Introduction

Lung cancer is one of the most common and leading causes of cancer deaths worldwide, with NSCLC (non-small cell lung cancer) being the most prevalent subtype.[1] Metastasis, the primary contributor to lung cancer mortality due to its ability to spread the disease systemically, has become a major focus of interest, particularly in the application of bioinformatics and artificial intelligence (AI).[2] From 2014 to 2018, the annual decline in lung cancer mortality in the U.S. accelerated to 5.5% in men and 4.4% in women, driven largely by advances in treatment for non-small cell lung cancer (NSCLC), whose two-year relative survival rate improved from 34% (2009–2010) to 42% (2015–2016). Despite these gains, NSCLC still accounts for the majority of lung cancer-related deaths.[3,4] The Indian subcontinent has a lower incidence rate of about 6.6 per 100,000 between 1990 and 2016 (10.3 to 11.2 in men and from 2.6 to 4.5 in women).[5] Lung cancer metastasis involves complex molecular signaling, particularly through protein–protein interactions that drive tumor spread.[6] Understanding these interactions is critical for identifying novel therapeutic targets and early biomarkers. Natural compounds have shown promise in inhibiting metastatic proteins but require computational modelling for effective screening.[2,7]

Although previous studies have mapped critical oncogenic pathways, predictive models integrating AI and PPI (protein–protein interaction) data remain limited. While traditional methods struggle to handle the high complexity and diversity of cancer metastasis, AI can significantly enhance the prediction of key metastatic drivers and their inhibition through *in silico* screening.[8] This integration can significantly bolster precision medicine efforts, with notable benefits for personalized cancer therapy. Major initiatives such as the TCGA (The Cancer Genome Atlas) and the Human Protein Atlas are attempting to investigate and better understand the molecular-level interactions behind metastasis.[9] Currently, the USA is the global leader in the development of AI-based mathematical modelling (use of algorithms to simulate biological behaviors and predict outcomes in complex systems) and drug discovery platforms to identify novel methods to target metastasis.[3]

Seven national bodies like ICMR, DBT, and CSIR are funding projects integrating bioinformatics and network biology (a systems biology approach to study interactions between biological entities) to combat cancer.[10] Furthermore, India has rich biodiversity, leading to increased interest in natural compound inhibition (the process of blocking disease-related proteins using naturally derived chemical compounds) as an anti-cancer drug.[11] AI and ML (machine learning) tools can swiftly analyze vast biological and medical datasets to predict protein functions, interactions, identify drug targets, and simulate molecular interactions.[8,12] In cancer research, AI models can detect hidden patterns in multi-omics data to forecast metastatic potential.[13] In the U.S. and India, AI is becoming central in drug repurposing, natural compound screening, and oncology diagnostics, with ML algorithms such as random forests, support vector machines, and deep learning models being used for PPI prediction.[14]

DOI: 10.36838/IJHSR88.90

Natural compound studies often rely on wet-lab approaches, which are time and resource-consuming, with computational approaches being generally underutilized.[15] Collaborations between AI developers and life scientists are still limited, and most research focuses on individual omics layers, rarely integrating multi-omics and network data.[16] Thus, it was hypothesized that AI-driven mathematical modelling, when integrated with network biology, can accurately predict PPIs driving NSCLC metastasis. Additionally, the computational screening of natural compound libraries can identify potential natural compound metastasis inhibitors, aiding drug development. The main objective of this paper is to contribute to the early prediction and prevention of metastasis in NSCLC patients, enable faster drug discovery by identifying bioactive natural compounds through AI models, and support the usage of indigenous medicinal compounds for cancer treatment, while contributing globally to the evolving AI-oncology intersection.

Methods

Data Collection and Pre-processing:

A multi-stage data collection and filtering process was used to identify robust gene candidates involved in NSCLC, integrating transcriptomic datasets from multiple sources and taking rigorous pre-processing steps. Differentially expressed genes were initially obtained from a TCGA-based dataset available on the Lung Cancer Explorer (LCE) portal. Both of the subtypes of LUAD (lung adenocarcinoma) and LUSC (lung squamous cell carcinoma) were taken into account during the analysis to ensure comprehensive subtype coverage. Non-coding and microRNA entries were filtered from the original dataset, with only protein-coding transcripts of likely functional significance retained. An adjusted p-value threshold of < 0.05 was chosen to pick up statistically significant genes, with a standardized mean difference (SMD) threshold of > 0.5 chosen to secure practical biological significance. SMD is a measure of effect size that expresses the difference between two group means relative to the variability (standard deviation) of the data, allowing comparison across studies with different measurement scales. Additionally, only genes were retained that were directionally consistent in both LUAD and LUSC, displaying uniform upregulation or uniform downregulation. To complement this data, expression and clinical metadata for LUAD and LUSC were retrieved from The Cancer Genome Atlas (TCGA) via the cBioportal platform. Only diploid samples, defined as exhibiting a regular genomic copy number without significant amplification or deletion, were selected for analysis. The dataset included z-score normalized mRNA expression data for both LUAD and LUSC diploid samples. Only those showing consistent upregulation or downregulation in both LUAD and LUSC subtypes were kept.

To further validate and emphasize relevance to cancer metastasis, three additional datasets were selected from the Gene Expression Omnibus (GEO): GSE161116, GSE166720, and GSE263726. These datasets were chosen based on their association with metastatic character in NSCLC. Specifically, they were prioritized because they included clear metastatic versus non-metastatic annotations (or metastatic stage grouping) and sufficient sample sizes to support reliable differential expression analysis. The DEG (differentially expressed gene(s)) lists for each of these datasets were generated using GEO2R, an in-built tool within GEO that implements standard statistical procedures to identify differentially expressed genes between two or more experimental groups. Default thresholds applied by GEO2R were used for the analysis, including a p-value cut-off of less than 0.05. The DEGs identified from each of the three studies were combined into a unified GEO-sourced DEG list. The union of the GEO-derived DEGs was intersected with the curated LCE–cBioPortal gene list to produce a final gene list.

Protein-Protein Interaction (PPI) Network Construction and Pathway Analysis:

To evaluate the structural relationships between metastasis-associated genes in NSCLC, a PPI network was constructed. The final list of 850 DEGs was queried against the STRING database using a custom Python script that implemented batch submission of gene symbols. The script parsed STRING's API and filtered results to include only interactions with a combined score ≥ 0.7 and limited to Homo sapiens. The resulting interaction table (in TSV 'tab-separated values' format) included columns for interacting protein pairs, interaction scores, and evidence types. This interaction matrix was imported into the Cytoscape software. Within Cytoscape, the Network Analyzer plugin computed multiple topological parameters. To gain functional insight into the metastasis-associated genes identified through expression integration and network modeling, pathway enrichment analysis was conducted using the Reactome Pathway Database. Statistical significance was determined using hypergeometric testing with Benjamini–Hochberg correction to control the false discovery rate (FDR).

Dynamic Modelling of PPI Networks:

While centrality highlights structurally important hubs, dynamic modeling evaluates how these genes behave over time under regulatory interactions, helping distinguish consistently influential drivers from nodes that are merely well connected. To simulate dynamic regulatory behavior within this network, a non-linear ODE model was implemented to capture the time evolution of protein concentrations across all the genes. A non-linear ordinary differential equation is an ODE in which the unknown function or its derivatives appear with powers, products, or other non-linear functions, so the equation cannot be written as a sum of terms where the function and its derivatives are only to the first power multiplied by functions of the independent variable. Each gene was modelled by the nonlinear equation $dx/dt = \alpha - \beta x + A \cdot x^2$. Both α and β were fixed at 1 across all genes, simplifying the dynamics while retaining general biochemical interpretability. Initial activity levels for each gene were randomly initialized within the range of 0.01 to 0.1 to mimic low resting expression. Simulations were conducted over 100 discrete time steps using the LSODA algorithm through SciPy's solve_ivp function, which adaptively

handles stiff and non-stiff systems. All simulation parameters, including step size and tolerances, were held constant across the entire gene set. The model produced a time-series matrix of gene expression trajectories. All the genes were retained throughout the simulation, and their dynamic profiles were ranked based on convergence behavior.

Virtual Screening of BUB1:

Molecular docking was performed to explore therapeutic inhibition of the key metastasis-associated gene *BUB1*. An *in silico* pipeline based on PyRx, AutoDock Tools, and Open Babel was used for the analysis. The 3D crystal structure of BUB1 was obtained from the RCSB Protein Data Bank in PDB (Protein Data Bank) format. The structure was pre-processed before docking using AutoDock Tools (ADT). This included: the removal of all water molecules, the addition of polar hydrogens to support accurate hydrogen bonding interactions, and the assignment of Kollman charges to all atoms. The following structure was then converted into PDBQT (AutoDock Protein Data Bank format with partial charges and atom types) format. For ligand preparation, a library of natural compounds was sourced from PubChem in SDF (Structure Data File) format and imported into PyRx. Using the Open Babel module integrated within PyRx, the compounds were converted into 3D conformers and energy-minimized using the MMFF94 force field to optimize their geometry. These processed ligands were then exported into PDBQT format, which is required for docking. A blind docking approach was employed to scan the entire surface of the BUB1 protein for potential binding sites. Within PyRx, the docking grid box was configured to encompass the full macromolecule to ensure that no potential binding cavity was overlooked. The center and dimensions of the grid box were manually set based on visual inspection of the protein's structure. Docking grid parameters were optimized once for the BUB1 binding site and then reused unchanged across all ligands to ensure consistent sampling and reproducibility. Docking was conducted using AutoDock Vina through the Vina Wizard interface in PyRx. Flexible ligand docking was performed against the prepared BUB1 protein, with AutoDock Vina applying its scoring algorithm to estimate binding affinities. The results were reported as binding free energies (ΔG, in kcal/mol), where more negative values indicated stronger predicted binding interactions. Following molecular docking, the top five ligands ranked by binding affinity were selected for detailed interaction analysis using BIOVIA Discovery Studio Visualizer. Upon opening each complex in Discovery Studio, the protein and ligand chains were visually inspected. For each loaded complex, the built-in "Receptor-Ligand Interactions" protocol was executed.

Machine Learning Model:

A supervised machine learning approach was implemented to predict the binding potential of natural compounds targeting the BUB1 protein. These compounds had previously been screened in the molecular docking pipeline and were selected from a curated natural product library retrieved from PubChem. The physicochemical descriptors of the ligands, such as molecular weight, topological polar surface area (TPSA), XLogP, the number of hydrogen bond donors and acceptors, rotatable bonds, and several additional drug-likeness indicators, were used as the feature set. TPSA is the sum of the surface areas of all polar atoms (typically oxygen and nitrogen) and their attached hydrogens in a molecule's 2D structure, used as a predictor of permeability and drug absorption. All descriptor data were compiled into a single feature matrix and paired with binary classification labels based on binding energy values derived from AutoDock Vina simulations. Compounds showing binding energies of –7.0 kcal/mol or better were labelled as 'active' and assigned a value of 1. Those with less favorable binding affinities were labelled 0. This threshold was selected to distinguish strong binders from weaker candidates while keeping the classification criterion both interpretable and stringent. The dataset was split into training and testing sets using an 80:20 ratio, with stratified sampling to maintain class distribution. Model performance was evaluated on the held-out test set using accuracy and ROC-AUC (with precision, recall, and F1-score reported as supporting metrics) to verify that predictive performance generalized beyond the training data. To interpret the influence of individual features on model output, SHAP (SHapley Additive exPlanations) values were computed using TreeExplainer. The SHAP summary plot highlighted the key features. In addition to binary class predictions, the model also generated probability scores for each compound using the predict_proba function. These confidence scores indicated the model's certainty in classifying a ligand as active and were subsequently used to prioritize candidates for downstream consideration. A final ranked list containing compound identifiers, predicted probabilities, and classification labels was exported for use in compound selection workflows.

Results and Discussion

Identification of Differentially Expressed Genes Across NSCLC Subtypes:

From the TCGA datasets, genes were considered differentially expressed based on standardized expression differences computed across metastatic and non-metastatic samples. In lung adenocarcinoma (LUAD), a total of 11,692 genes were differentially expressed, of which 7,066 were upregulated, and 4,626 were downregulated. Similarly, analysis of the lung squamous cell carcinoma (LUSC) dataset identified 12,039 differentially expressed genes (DEGs), including 6,799 upregulated and 5,240 downregulated. The intersection yielded 6,032 genes that were consistently upregulated or downregulated across both NSCLC subtypes. After performing DE analysis for chosen IDs, GSE161116, GSE166720, and GSE263726, a separate number of DEGs were identified for each. These GSE IDs were selected based on the presence of well-annotated metastatic phenotypes and control groups. GSE161116 returned 352 DEGs, 342 of which were downregulated. GSE166720 and GSE263726 identified 13,374 and 7,819 DEGs, respectively, with distributions that included both upregulated and downregulated candidates (Figure 1).

DOI: 10.36838/IJHSR88.90

After consolidation and removal of duplicates, the GEO-derived DEG pool comprised 16,861 unique genes. To establish a final metastasis-associated signature, the 6,032 subtype-consistent TCGA genes were intersected with the GEO-derived DEG pool. This cross-validation step resulted in a core set of 858 genes that were reproducibly dysregulated in both large-scale RNA-seq and independent microarray datasets. These 858 genes represent highly statistically robust and biologically validated candidates associated with NSCLC metastasis.

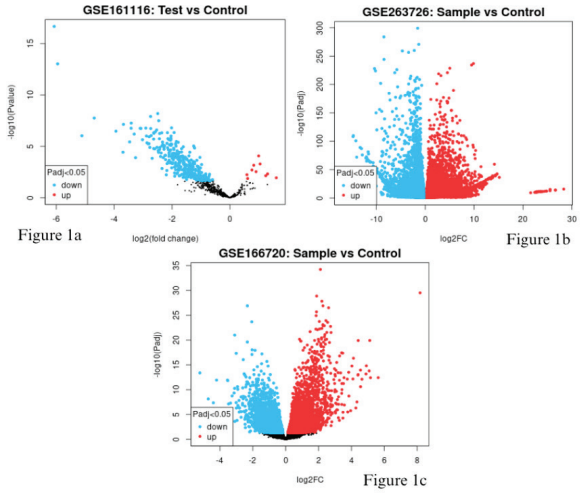


Figure 1: Differentially expressed genes in NSCLC metastasis studies from GEO (GSE161116, GSE166720, GSE263726). Volcano plots showing differential gene expression across the three GEO datasets. Each point represents a gene, plotted by $\log_2$ fold change (x-axis) and $-\log_{10}$ adjusted p-value (y-axis). Genes with statistically significant differential expression (adj p-value < 0.05) are colored: upregulated genes in red, downregulated genes in blue, and non-significant genes in black. These plots highlight the distribution and magnitude of transcriptional dysregulation between control and metastatic lung cancer samples in each study.

Network Construction and Hub Gene Identification:

The functional relationships among the 858 GEO-validated genes implicated in NSCLC metastasis were mapped in the form of a protein–protein interaction (PPI) network. The resulting network consisted of 858 nodes and 5,321 undirected edges, indicating a densely interconnected regulatory architecture (Figure 2). Topological features, including degree, betweenness, closeness, and clustering coefficient, were computed for each gene. Degree centrality was prioritized as the primary measure of node importance due to its intuitive representation of connectivity. Based on all these centrality metrics, the ten most highly connected hub genes were identified (Table 1). Among these, *BRCA1* showed the highest degree with 67 direct connections, marking it as a key regulatory node within the network.

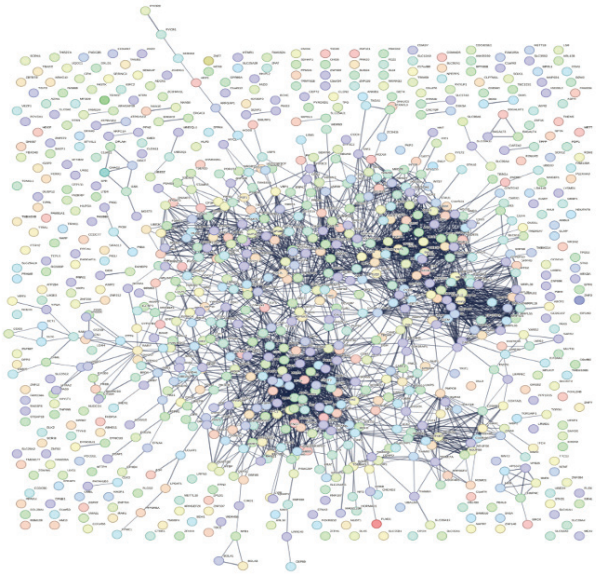


Figure 2: A Protein–Protein Interaction (PPI) Network of DEGs derived from TCGA (LUAD and LUSC) and GEO (GSE161116, GSE166720, GSE263726). The full PPI network was constructed using the STRING database for the 858 GEO-validated, TCGA-consistent differentially expressed genes (858 nodes; 5,321 edges). Round nodes represent proteins and edges indicate high-confidence interactions (combined score > 0.7).

Table 1: Summary of Top 10 Hub Genes Identified from Static Centrality Analysis of the NSCLC Metastasis Network. The table lists the top 10 hub genes ranked by degree centrality in the STRING-derived protein–protein interaction (PPI) network of 858 GEO-validated metastasis-associated genes.

S. No.	Gene	Degree	Betweenness Centrality	Closeness Centrality	Clustering Coefficient
1	*BRCA1*	67	0.1144026131	0.3372147182	0.1203075531
2	*EXO1*	57	0.01958042675	0.3070398643	0.2675438596
3	*POLR1B*	50	0.06076667154	0.3265674335	0.2563265306
4	*MRPS12*	48	0.01618940823	0.3116659492	0.4078014184
5	*NOP56*	47	0.02127416617	0.3097988875	0.3977798335
6	*FBL*	46	0.03830293692	0.3186619718	0.3806763285
7	*BYSL*	45	0.0215462708	0.2965997542	0.4141414141
8	*BUB1*	45	0.02350532586	0.3123382226	0.3696969697
9	*TOP2A*	45	0.01876130807	0.312203536	0.3808080808
10	*CDC45*	45	0.01509178348	0.3058724123	0.3707070707

Dynamic Modeling of Network Behavior:

To model the dynamic regulatory behavior of metastasis-associated genes in NSCLC, a nonlinear ordinary differential equation (ODE) simulation was employed using the STRING-derived protein–protein interaction (PPI) network. The validated list of 858 genes was mapped onto the network, and the corresponding adjacency matrix was extracted. Each node represented a protein and was modelled according to the equation $dx/dt = \alpha - \beta x + A \cdot x^2$, where x is the protein activity level, α is the synthesis rate, β is the degradation rate, and A is the weighted adjacency matrix. Initial conditions were randomly sampled between 0.01 and 0.1 to reflect low basal activity. Simulations were conducted for 50 time steps using the LSODA solver in SciPy's solve_ivp module. Nodes exhibiting oscillatory or null behavior were filtered out. A node was considered oscillatory if the time-series profile exhibited multiple

peaks or fluctuations without convergence, with a standard deviation > 0.1 over the final 50% of the simulation period. Final trajectories were compiled into a time-series CSV (comma-separated values) matrix, with temporal plots generated for select genes (Figure 3). All ten of the top-ranked genes from the static centrality analysis–*BRCA1, EXO1, TOP2A, NOP56, MRPS12, BUB1, CDC45, RRM2, NOP58,* and *FBL*–exhibited stable, non-zero equilibrium states. These profiles showed initial exponential increases in activity followed by smooth convergence to plateaus, typically within 10–15 time steps. To assess coherence between topological and kinetic importance, the top 10 dynamic nodes with the top 10 from static centrality measures were compared. Nine genes were shared: *BRCA1, BUB1, CDC45, EXO1, FBL, MRPS12, NOP56, NOP58,* and *TOP2A*. The only exception was BYSL, which ranked just outside the top 10 dynamically at position 11. When extended to the top 30 genes, 29 overlapped, demonstrating strong convergence between structural prominence and temporal persistence (Figure 4). This alignment supports the biological importance of the selected genes and affirms their prioritization for downstream experimental or therapeutic targeting.

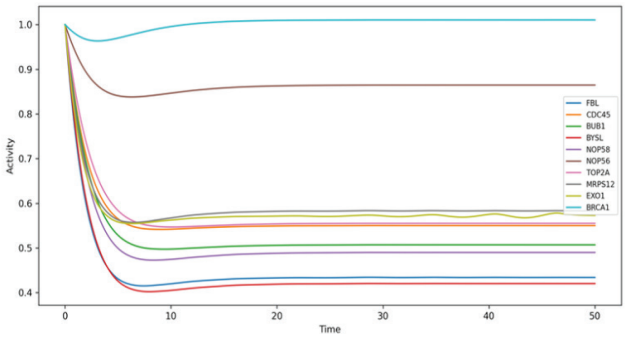


Figure 3: Dynamic Activity Profiles of Top 10 Hub Genes in Simulated PPI Network derived from TCGA (LUAD and LUSC) and GEO (GSE161116, GSE166720, GSE263726). Time-course trajectories depicting simulated activity dynamics of the top 10 hub genes (*BRCA1, EXO1, TOP2A, NOP56, MRPS12, BUB1, CDC45, RRM2, NOP58, FBL*) within the dynamic protein–protein interaction (PPI) network. Each curve represents the expression activity of a single gene over 50 simulation time steps based on a nonlinear ODE system. All trajectories converge to non-zero steady states.

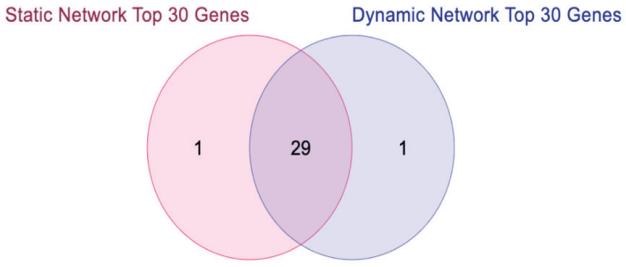


Figure 4: Comparison of Top-Ranked Genes Between Static and Dynamic Network Models derived from TCGA (LUAD and LUSC) and GEO (GSE161116, GSE166720, GSE263726). Venn diagram illustrating the overlap between the top 30 genes ranked by static degree centrality (left, red) and those prioritized through dynamic modeling of the same 858-gene protein–protein interaction network (right, blue). Of the 30 highest-ranked genes from each model, 29 were shared, demonstrating a high degree of agreement between structural centrality and dynamic stability-based prioritization. Only one gene was unique to each ranking list, validating the robustness of the consensus-derived gene set.

Pathway Enrichment Analysis:

After performing pathway analysis, a total of 107 pathways were significantly enriched (FDR < 0.05), with the majority clustering around cell cycle regulation, DNA replication, and checkpoint transitions. The top five enriched pathways identified were Mitotic G1 phase and G1/S transition, Cell Cycle, Mitotic, G1/S Transition, and Cell Cycle Checkpoints. These pathways collectively cover a substantial portion of the functional landscape defined by the 858-gene panel. Their prominence suggests a tightly regulated disruption of cell cycle progression and genomic maintenance in NSCLC metastasis. Many of the top-ranking hub genes from the network analysis, including *BRCA1, BUB1, TOP2A, CDC45*, and *EXO1*, were also present in these enriched pathways. Their recurrence across static centrality, dynamic simulation, and functional annotation further validates their mechanistic relevance.

Virtual Screening and Docking Analysis:

To identify candidate natural compounds capable of targeting key metastatic regulators in NSCLC, molecular docking was conducted for the BUB1 protein using a structure-based virtual screening approach. BUB1 was selected as the docking target due to its high topological centrality, temporal stability in dynamic modeling, and consistent enrichment in mitosis-related pathways. A total of 1,869 natural product-derived ligands were compiled from PubChem and prepared using Open Babel within the PyRx interface. AutoDock Vina returned binding free energy estimates (ΔG, in kcal/mol) for each ligand–protein complex. All results were aggregated and ranked by ascending ΔG values, and top compounds were identified (Table 2). The top binding affinities ranged from −11.7 to −9.6 kcal/mol, indicating strong interaction potential. The top-ranked compounds demonstrated consistent docking across multiple simulation runs and structural replicates. CID_135398501, the top hit, exhibited a predicted binding energy of −11.8 kcal/mol, placing it well within the range associated with high-affinity ligand binding. The majority of the top ligands contained heterocyclic cores and polar functional groups conducive to hydrogen bonding and hydrophobic pocket fitting. Structural analysis using PyMOL revealed favorable interactions within BUB1's ATP-binding domain and adjacent regulatory sites. Three of the five leading compounds –CID 135398501 (AKT inhibitor VIII), CID 11751922 (CDK inhibitor AT-7519), and CID 5154 (sanguinarine) are documented in PubChem as kinase inhibitors or bioactive antitumor agents, lending support to their pharmacophore relevance. The remaining hits show anticancer activity but lack specific kinase annotation. BIOVIA Discovery Studio provided a 2D interaction diagram for the top 5 ligands (Figure 5) showing the types of interactions.

DOI: 10.36838/IJHSR88.90

Table 2: Top Five Ligands Identified from Virtual Screening Against the BUB1 protein in NSCLC. List of top-performing ligands from structure-based virtual screening against the BUB1 protein using AutoDock Vina. Compounds were ranked based on binding affinity (ΔG, kcal/mol). Some of these compounds are known or suspected kinase inhibitors, suggesting therapeutic relevance in NSCLC metastasis inhibition.

S.No.	Rank	Compound Name / ID	Binding Energy (kcal/mol)
1	1	CID_135398501	−11.8
2	2	CID_5154	−11.7
3	3	CID_11751922	−11.5
4	4	CID_24882195	−11.5
5	5	CID_5281600	−11.4

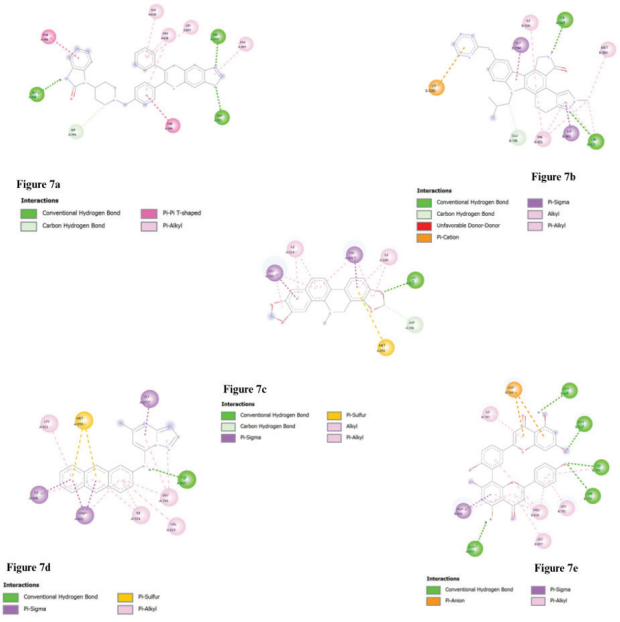


Figure 5: 2D representation of docking using Biovia Discovery Studio top poses after virtual screening. 2D interaction diagrams of the top five high-affinity ligands (pose 1) docked to the BUB1 protein (CID_135398501, CID_5154, CID_11751922, CID_24882195, CID_5281600). Visualizations were generated using BIOVIA Discovery Studio based on AutoDock Vina output poses.

Machine learning model:

To evaluate the drug-likeliness and binding potential of the screened natural compound library, a machine learning–based classification framework was employed to predict ligand activity against the BUB1 protein with two models, XGBoost and Random Forest. With both demonstrating acceptable predictive behavior, XGBoost was chosen for feature importance analysis and compound prioritization due to its interpretability via SHAP and efficiency in handling tabular data. Feature importance was explored using SHAP (SHapley Additive exPlanations), a technique that offers insight into how each input feature contributes to the final prediction. The SHAP summary plot showed that XLogP, topological polar surface area (TPSA), and the number of rotatable bonds had the greatest impact on classification outcomes. These descriptors, which reflect molecular hydrophobicity, polarity, and flexibility, are commonly linked to drug-like behavior. The model relied on these variables to distinguish between strong and weak binders. A global summary of these findings is presented in Figure 6, offering a clear visualization of overall feature relevance. Once the model was trained, it was used to assign binding activity probabilities to every compound in the dataset. Compounds predicted as active (based on classification probabilities) were flagged for prioritization. A total of 367 molecules met this criterion and were advanced for downstream consideration. The original training labels were based on docking scores below −7.0 kcal/mol, but the final model predictions relied solely on descriptor patterns. These selected compounds represented a wide range of chemical scaffolds and exhibited high confidence scores, making them strong candidates for redocking or experimental validation.

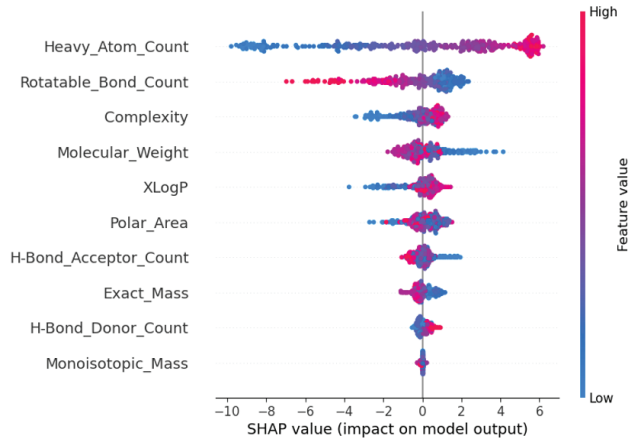


Figure 6: SHAP Summary Plot of Feature Importance in the XGBoost Model trained on the docked natural compound set (n = 1,869 ligands; "active" label defined as AutoDock Vina binding energy ≤ −7.0 kcal/mol). SHAP (SHapley Additive exPlanations) values summarize the contribution of each molecular descriptor to the classification of compounds as active or inactive against BUB1. Each dot represents a SHAP value for a given compound and feature. The horizontal axis indicates the impact on the model output, while color represents the feature value (red = high, blue = low).

This study presents a robust computational framework with multiple layers for identifying metastasis-associated gene signatures in non-small cell lung cancer (NSCLC). Merging transcriptomic data from both The Cancer Genome Atlas (TCGA) datasets and Gene Expression Omnibus (GEO), followed by protein–protein interaction (PPI) network construction, dynamical modeling, and virtual compound screening, enabled the identification of a high-confidence list of 858 genes with metastasis signature. Unlike previous studies, which rely on single-cohort or platform-specific analyses, this combined approach takes advantage of multi-cohort and cross-platform validation. This method arguably strengthens both biological relevance and further utility. The transcriptomic integration phase served to refine the candidate gene list by ensuring consistency across NSCLC subtypes and platforms. The initial intersection of LUAD and LUSC transcriptomic datasets excluded subtype-specific variability, reinforcing the generalizability of the metastasis-associated signature across NSCLC. This strategy was further strengthened through validation using GEO-derived gene expression data. By lever-

aging datasets across distinct platforms and clinical contexts, the resulting gene set exhibits strong cross-platform robustness. Such integrative approaches are increasingly recognized for their biomarker discovery power. For example, a research group constructed a multi-gene LUAD prognostic model by integrating TCGA and GEO cohorts, highlighting how transcriptomic convergence improves generalizability and clinical utility.[17] While our approach shares this integrative nature, it differs by highlighting metastasis-specific signatures.

The PPI network of the 858-gene set revealed a scale-free topology enriched in cell cycle and DNA repair regulators. Among these, some genes (e.g., *BRCA1, BUB1*, and *TOP2A*) stood out with high degree centrality and were further validated by being prioritized through nonlinear ordinary differential equation (ODE) modeling. These genes maintained stable, non-zero equilibrium states over simulated timesteps, suggesting that they play a persistent regulatory role in maintaining network homeostasis under dynamic conditions relevant to metastasis. This points toward a form of kinetic resilience (the ability of a node to retain regulatory influence over time despite perturbations) that may explain their influence in NSCLC progression. This interpretation aligns with Ma *et al*., who showed *BRCA1*'s involvement in homologous recombination repair (HRR) and its predictive value in NSCLC prognosis.[18] Additionally, systems-level network analysis was conducted by another study, where they emphasized how central mitotic regulators, like *BUB1* and *TOP2A*, often correspond to key phenotypic drivers in cancer.[19] These studies validate our reliance on static centrality as a biologically meaningful criterion.

To contextualize these hubs functionally, enrichment analysis using Reactome revealed that the identified hub genes were involved in G1/S checkpoint control, mitotic spindle formation, and chromosomal segregation, to name a few. These findings suggest that dysregulation of mitotic fidelity may be a significant, though not exclusive, contributor to NSCLC metastatic potential. A possible explanation could be that compromised checkpoint control permits chromosomal instability, facilitating malignant transformation. This interpretation finds some support in a study that demonstrated that alterations in oxidative stress response and DNA damage repair pathways stratify LUAD subtypes with poor prognosis.[20] Additionally, enrichment of RNA splicing, mitochondrial translation, and proteasome pathways aligns with the findings of a study that identified similar biological programs in LUSC cohorts associated with immune infiltration and poor survival.[21]

With this mechanistic backdrop established, virtual screening revealed several candidate compounds with strong predicted binding affinities toward BUB1, including the known kinase inhibitors sanguinarine and AT-7519. These findings are promising given the existing evidence of *BUB1* overexpression in NSCLC and its role in tumor proliferation. Although direct inhibitors of BUB1 remain limited in clinical use, several BUB1-targeted compounds are currently under preclinical evaluation for their potential to disrupt mitotic fidelity and tumor growth. Related kinase inhibitors have shown efficacy in preclinical studies. This supports the potential of BUB1-targeted compounds as viable anticancer agents. It remains to be tested, however, whether such ligands retain efficacy in complex in vivo tumor microenvironments. The identification of phytochemical ligands like sanguinarine also aligns with prior efforts exploring phytochemical inhibition of mitotic regulators, although direct targeting of BUB1 remains unverified.

To distill these docking outputs into a prioritized list, an XGBoost classifier was trained on molecular descriptors and refined using SHAP value interpretation. The model highlighted XLogP, TPSA, and rotatable bonds as key predictors of binding affinity, confirming established structure–activity relationships in medicinal chemistry. Pashaei *et al*. similarly demonstrated the efficacy of machine learning coupled with network-based analysis for high-throughput drug screening and compound interpretation. While our model exhibited robust feature interpretability, its predictions are necessarily constrained by the resolution of available docking scores. Of the 1,869 compounds screened, 367 were prioritized as high-confidence candidates, providing a tractable shortlist for future validation. Future applications should include experimental validation of predicted targets as well as ligands, for instance, with CRISPR-mediated gene knockout for functional characterization or SPR assay for confirmation of compound binding affinity.

This study acknowledges some limitations that it is primarily computational and relies on *in silico* predictions, so results should be interpreted as hypothesis-generating rather than definitive evidence of biological causality. No wet lab validation was performed, meaning the proposed role of BUB1 and the prioritized natural compounds require experimental confirmation (for example, cell-based assays of migration or invasion and target engagement). In addition, docking scores and the chosen activity threshold are approximations that depend on the docking setup and scoring function, so binding affinity rankings may shift under alternative parameterizations or more rigorous free energy methods.

Conclusion

Overall, this work demonstrates the power of merging transcriptomic integration, network biology, dynamical systems models, and AI-improved virtual screening to identify regulatory hubs alongside therapeutic leads for NSCLC metastasis. Concurrence of biologically reasonable conclusions for an extended set of analytical spaces demonstrates internal validity of the analytical pipeline. Such multidimensional concurrence reinforces confidence in prioritized targets alongside their translational validity. At a high level, our systems-level approach potentially offers a starting point for stabilization of biomarkers alongside target fidelity improvement for molecularly heterogeneous cancers like NSCLC, wherein traditional single-layer analyses are suboptimal.

Acknowledgments

Prachet Patakula would like to acknowledge Aashna Saraf, Founder of CreatED, for providing valuable feedback and guidance throughout the project. I would also like to acknowledge my parents for their unwavering support and motivation throughout.

DOI: 10.36838/IJHSR88.90

References

1. Mehta, A. A.; Pavithran, K.; Nair, P. K.; Vazhoor, V.; Gutjahr, G.; Lakshmi Priya, V. P., Epidemiological and histopathological profile of lung Cancer: Insights from a 15-year cross-sectional study at a tertiary care centre in South India. *Global Epidemiology* **2025**, *9*, 100208.
2. Liao, J.; Li, X.; Gan, Y.; Han, S.; Rong, P.; Wang, W.; Li, W.; Zhou, L., Artificial intelligence assists precision medicine in cancer treatment. *Frontiers in Oncology* **2023**, *Volume 12 - 2022*.
3. Weaver, I. N.; Weaver, D. F., Drug design and discovery: translational biomedical science varies among countries. *Clin Transl Sci* **2013**, *6* (5), 409-13.
4. Siegel, R. L.; Miller, K. D.; Fuchs, H. E.; Jemal, A., Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians* **2021**, *71* (1), 7-33.
5. Collaborators, I. S.-L. D. B. I. C., The burden of cancers and their variations across the states of India: the Global Burden of Disease Study 1990-2016. *Lancet Oncol* **2018**, *19* (10), 1289-1306.
6. Jamil, A.; Kasi, A., Lung Metastasis. In StatPearls, StatPearls Publishing Copyright © 2025, StatPearls Publishing LLC.: Treasure Island (FL) ineligible companies. Disclosure: Anup Kasi declares no relevant financial relationships with ineligible companies. 2025.
7. Chunarkar-Patil, P.; Kaleem, M.; Mishra, R.; Ray, S.; Ahmad, A.; Verma, D.; Bhayye, S.; Dubey, R.; Singh, H. N.; Kumar, S., Anticancer Drug Discovery Based on Natural Products: From Computational Approaches to Clinical Studies. *Biomedicines* **2024**, *12* (1).
8. Bi, W. L.; Hosny, A.; Schabath, M. B.; Giger, M. L.; Birkbak, N. J.; Mehrtash, A.; Allison, T.; Arnaout, O.; Abbosh, C.; Dunn, I. F.; Mak, R. H.; Tamimi, R. M.; Tempany, C. M.; Swanton, C.; Hoffmann, U.; Schwartz, L. H.; Gillies, R. J.; Huang, R. Y.; Aerts, H., Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* **2019**, *69* (2), 127-157.
9. Tomczak, K.; Czerwińska, P.; Wiznerowicz, M., The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* **2015**, *19* (1a), A68-77.
10. Agarwal, A.; Agrawal, P.; Sharma, A.; Kumar, V.; Mugdal, C.; Dhall, A.; Raghava, G. P. S., A repository of web-based bioinformatics resources developed in India. *bioRxiv* **2020**, 2020.01.21.855627.
11. Siddiqui, A. J.; Jahan, S.; Singh, R.; Saxena, J.; Ashraf, S. A.; Khan, A.; Choudhary, R. K.; Balakrishnan, S.; Badraoui, R.; Bardakci, F.; Adnan, M., Plants in Anticancer Drug Discovery: From Molecular Mechanism to Chemoprevention. *Biomed Res Int* **2022**, *2022* (1), 5425485.
12. Paul, D.; Sanap, G.; Shenoy, S.; Kalyane, D.; Kalia, K.; Tekade, R. K., Artificial intelligence in drug discovery and development. *Drug Discov Today* **2021**, *26* (1), 80-93.
13. Bhinder, B.; Gilvary, C.; Madhukar, N. S.; Elemento, O., Artificial Intelligence in Cancer Research and Precision Medicine. *Cancer Discov* **2021**, *11* (4), 900-915.
14. Serrano, D. R.; Luciano, F. C.; Anaya, B. J.; Ongoren, B.; Kara, A.; Molina, G.; Ramirez, B. I.; Sánchez-Guirales, S. A.; Simon, J. A.; Tomietto, G.; Rapti, C.; Ruiz, H. K.; Rawat, S.; Kumar, D.; Lalatsa, A., Artificial Intelligence (AI) Applications in Drug Discovery and Drug Delivery: Revolutionizing Personalized Medicine. *Pharmaceutics* **2024**, *16* (10).
15. Medina-Franco, J. L., Computational Approaches for the Discovery and Development of Pharmacologically Active Natural Products. *Biomolecules* **2021**, *11* (5).
16. Chen, C.; Wang, J.; Pan, D.; Wang, X.; Xu, Y.; Yan, J.; Wang, L.; Yang, X.; Yang, M.; Liu, G. P., Applications of multi-omics analysis in human diseases. *MedComm (2020)* **2023**, *4* (4), e315.
17. Zheng, W.; Zhou, C.; Xue, Z.; Qiao, L.; Wang, J.; Lu, F., Integrative analysis of a novel signature incorporating metabolism and stemness-related genes for risk stratification and assessing clinical outcomes and therapeutic responses in lung adenocarcinoma. *BMC Cancer* **2025**, *25* (1), 591.
18. Ma, Y.; Huang, J.; He, L.; Du, J.; Liu, L.; Li, X.; Jiao, P.; Wu, X.; Zhou, W.; Xu, X.; Yang, L.; Di, J.; Zhu, C.; Li, L.; Liu, D.; Wang, Z., Evaluation of two algorithms measuring homologous recombination deficiency status in prognostic assessment for treatment-naïve non-small cell lung cancer. *Chin J Cancer Res* **2025**, *37* (3), 352-364.
19. Pashaei, E.; Liu, S.; Li, K.; Zang, Y.; Yang, L.; Lautenschlaeger, T.; Huang, J.; Lu, X.; Wan, J., DiCE: differential centrality-ensemble analysis based on gene expression profiles and protein-protein interaction network. *bioRxiv* **2025**.
20. Rao, W.; Zhang, Q.; Dai, X.; Yang, Y.; Lei, Z.; Kuang, X.; Xiao, H.; Zhu, J.; Xiong, Y.; Wang, D.; Yang, L., A three-subtype prognostic classification based on base excision repair and oxidative stress genes in lung adenocarcinoma and its relationship with tumor microenvironment. *Sci Rep* **2025**, *15* (1), 16647.
21. Chen, Y.; Wang, M., Revealing roles of PANoptosis-related genes in prognosis and molecular subtypes in lung squamous cell carcinoma by integrated bioinformatic analyses and experiments. *Clin Exp Med* **2025**, *25* (1), 154.

Authors

Patakula Patakula is presently a Grade 12 student at Manthan School, passionate about computational biology and mathematical modeling. His research on vaccine decision modelling using game theory has been recognized with the CREST Gold Award and IEOM publication honors. He plans to pursue bioengineering and medicine to advance translational research.

Nirupma Singh is a Bioinformatics Scientist with a doctorate in Biotechnology and Bioinformatics from the University of Delhi, with six years of hands-on research and development experience. Her journey is marked by a robust foundation in Machine Learning and Python, with five years of expertise. Proficient in Linux and cloud servers like AWS. She excels in structural biology, systems biology, protein/gene network analysis, data mining, and computational genomics.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Investigating Metabolic Stress *In Vivo*: Does Acute Glucose Deprivation Replicate Aspects of AD Pathogenesis?

Edward Yan

Westlake High School, 4100 Westbank Drive, Austin, TX, 78746, USA; ey6bbm@gmail.com

ABSTRACT: Alzheimer's disease (AD) is a debilitating neurodegenerative condition that predominantly affects memory and cognitive function. Proposed in 1992, the amyloid cascade hypothesis (ACH) states that the accumulation of amyloid-beta (Aβ) plaques in the brain is the primary driver of AD. However, numerous clinical trials targeting the molecular basis of the ACH have shown limited clinical efficacy. One emerging alternative hypothesis is that endoplasmic reticulum (ER) stress drives AD pathogenesis. This study aimed to investigate the relationship between ER stress and memory impairment *in vivo* using the fruit fly (*Drosophila melanogaster*). Mild and severe ER stress were induced via varying lengths of glucose deprivation (GD), and the T-maze test assessed effects on spatial memory after positive-reinforcement classical conditioning. Mild ER stress did not significantly impair spatial memory compared to controls, but surprisingly, neither did severe ER stress under the selected experimental conditions. Flies acclimated to the T-maze setup exhibited less average distance traveled away from baseline towards a positive reward stimulus following severe ER stress compared to non-acclimatized controls. These findings suggest a non-linear relationship between ER stress and cognitive function and contribute to growing evidence that ER stress may have a context-dependent role in AD development.

KEYWORDS: Animal Behavior, Neuroscience, Alzheimer's Disease (AD), Endoplasmic Reticulum (ER) Stress, "Type 3" Diabetes.

Introduction

In the United States (U.S.), approximately 6.9 million people aged 65 and older are affected by Alzheimer's disease (AD), a debilitating neurodegenerative disorder that primarily affects memory and motor coordination.[1] AD is the seventh leading cause of death in the U.S., killing nearly 120,000 individuals each year.[1] The consequences of AD extend beyond those afflicted, with an estimated 11 million Americans involved in the care of AD patients and 18.4 billion work hours (h) attributed to AD patient care in 2023.[1] The estimated total of Americans immediately impacted by AD is 17.9 million, or 5.3% of the population.[2] Males and females have a comparable incidence rate, despite females having a higher diagnosis rate due to having longer average lifespans.[3] In 2024, the total financial cost of AD in the U.S. was $360 billion for long-term health care and hospice services.[1] AD is becoming a growing demographic within the United States' increasingly aging population; between 2000 and 2021, the rate of AD deaths increased by 140% and the number of people living with AD in the U.S. is projected to double to 13.8 million by 2060.[1] It is therefore essential to understand the molecular factors responsible for AD onset and progression to suggest potential targets for treatment; such efforts have already yielded an extensive body of knowledge.

The majority of AD research refers to the amyloid cascade hypothesis (ACH).[4] The ACH asserts that oligomers of amyloid-beta (Aβ) and tau proteins render themselves neurotoxic to the human brain and progressively induce loss of motor function and cognitive decline.[4] Although this hypothesis has been generally accepted by the scientific community since 1992, clinical trials that have attempted to target AD through the molecular basis of this hypothesis, by directly targeting and destroying Aβ plaques or tau aggregates, have resulted in varying, inconsistent results. Levin *et al.* examined several clinical trials that targeted the ACH, finding that most had results that merely slowed down the progression of the disease as opposed to reversing or halting it.[5] For example, in a Phase III clinical trial, aducanumab, which acts by recognizing Aβ plaques and marking them for degradation by microglia, was found to slow down AD progression by 25%, but had no curative effect.[6] Further studies on Phase I clinical trials using donanemab and lecanemab, two other proposed AD drugs targeting the ACH, found "mixed or statistically insignificant results".[5,7,8] Although Levin *et al.* did not explicitly refute the ACH, they suggested limitations to its usefulness in informing therapeutic strategies for patients, claiming that the ACH may only "pertain to the preclinical [asymptomatic] stage of AD, prior to symptomatic onset".[5] Therefore, this suggests that there may be alternative or complementary hypotheses for AD pathogenesis that could serve as new avenues for therapeutic intervention.

Among alternative hypotheses that have recently gained prominence, it is suggested that metabolic dysregulation, particularly endoplasmic reticulum (ER) stress and the unfolded protein response (UPR), plays a significant role in AD pathogenesis.[9,10] ER stress occurs when the capacity of the ER to fold proteins is overloaded, and proteostasis is subsequently hindered due to an accumulation of unfolded or misfolded proteins in the ER.[11] The UPR is the physiological phenomenon that occurs when a cell responds to ER stress and attempts to mitigate it, occurring via the actions of three key enzymes: IRE1, ATF6, and PERK.[11] Central to the ER stress hypothesis

DOI: 10.36838/IJHSR88.98

is that the accumulation of Aβ and tau proteins causes a significant increase in ER stress and thus disrupts protein folding and, in turn, neuronal cell viability.[10]

The role that ER stress plays in AD pathogenesis is a topic of increasing interest and scientific scrutiny. Salminen *et al.* were the first to suggest a correlation between ER stress, the UPR, and AD pathogenesis in a prominent review article published in 2009.[12] Others followed to experimentally test whether ER stress plays a role in AD pathogenesis through the neurotoxic accumulation of Aβ and tau, upregulating ER stress and thereby causing cognitive decline.[9,10,12] A study in which rats were administered tunicamycin, an ER stress inducer, into their brain ventricles found a significant reduction in spatial memory, defined as memory associated with navigation and recall of spatial relationships, when rats were tasked with finding a hidden platform under the water's surface in a Morris water maze test.[13] Surprisingly, Hashimoto *et al.* determined that extracted cells from rat brains exposed to thapsigargin, another ER stress inducer, at concentrations of 2–5 μM, conferred no significant effects on spatial memory, thereby refuting their own initial hypothesis that ER stress would result in neuronal cell death.[14] A possible explanation for this is suggested in an earlier study by Fouillet *et al.*, which found that mild ER stress in mouse models, fruit fly models, and human neuroblastoma cells promoted neuroprotective autophagy, thereby inhibiting neuronal cell death through an adaptive response against apoptosis.[15]

There is growing evidence that glucose deprivation (GD) induces ER stress via the PERK pathway, akin to the extent observed for tunicamycin and thapsigargin.[19,20] The ability for GD to cause ER stress has been observed in a range of health conditions in humans, such as ischemic stroke and diabetes in addition to AD.[19,21] Particularly with regards to diabetes mellitus, AD has well-documented molecular and biochemical features overlapping with both type 1 and type 2 diabetes (T1D/T2D), and also represents a form of insulin resistance specifically involving the brain. Owing to this, AD has been increasingly referred to as "type 3 diabetes", or T3D.[22] This may extend to glucose dyshomeostasis in general; notably, García de la Cadena *et al.* observed that acute GD could upregulate apoptosis in critical memory-regulating cells such as neurons in the hippocampus.[19] As such, the question arises: does acute glucose deprivation *in vivo*, and the subsequent ER stress this induces, replicate aspects of AD pathogenesis in the form of reduced spatial memory?

This question would ultimately serve as central to my experimentation. I hypothesize that, consistent with Fouillet *et al.* and Lin *et al.*, mild ER stress induced by a short duration of GD, lasting 1 h, would not significantly impair the spatial memory of fruit flies (Drosophila melanogaster), but that heavy ER stress induced by a longer duration of GD, lasting 3 h, would lead to a decrease in spatial memory.[13,15] This study aims to bridge the gap between understanding whether varying degrees of ER stress affect spatial memory responses comparable to those observed in AD *in vivo*.

Methods

Materials and experimental setup:

Fruit flies (D. melanogaster, hereby referred to as "flies" interchangeably) were obtained from the Carolina Biological Supply Company (CBSC) in Burlington, NC, as living, wild-type adults with 25–30 flies per vial. Each trial featured a total of 12 vials for a cumulative sample size of 300–360 flies per trial. Across 5 trials, this yielded a cohort size of 1,500–1,800 flies. Two vials (50–60 flies) were arbitrarily selected via random number generator at RANDOM.ORG and assigned to one of six experimental groups. Adults were segregated from larvae and pupae, and were transferred into a cylindrical vial measuring 10 centimeters (cm) by 4 cm containing a medium formula of 2 grams (g) active dry yeast, 2 teaspoons (tsp) instant medium from CBSC, and 10 mililiters (mL) drinking water. Vials were obtained from CBSC, and yeast, water, and reagent-grade sucrose were obtained from Red Star, Crystal Geyser, and ALDON, respectively. Any larvae in the original shipping vial were transferred to a mini-freezer obtained from INSIGNIA™ and euthanized via freezing for 100 h, a well-documented euthanasia procedure.[23] Flies maintained an equal gender ratio, analogous to documented human incidence rates for AD.[3]

Two odorants were used: aloe vera and Moroccan rose, obtained in hydrosol form from George's and Whole Foods, respectively. Both odorants are harmless and were used to induce classical conditioning, or the association of a stimulus with a positive or negative outcome.[24] Flies were anesthetized with CO_2 generated from Alka-Seltzer® effervescent tablets prior to transfer to a housing tube with food medium (Figure 1). Flies were given 1–3 days (d) to acclimate pre-experimentation. Flies were segregated into 6 groups, as shown in Figure 2. Flies in groups with "NCC", or no classical conditioning, did not associate any odorant with a reward, whereas flies in groups labelled "YCC", or literally "yes classical conditioning", (Figure 2) received a reward when subjected to rose odorant, where the tube was coated with both rose hydrosol and a "sweet medium", composed as shown in Figure 3.

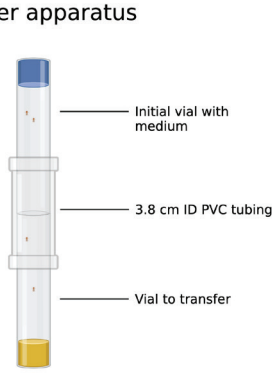


Figure 1: The transfer apparatus is used for transporting *D. melanogaster* between two vials. The transfer apparatus for the transfer of flies from their initial vials into a destination vial (vial to transfer) was as shown. This tube was tapped onto a table to knock flies into the vial to transfer, and the initial tube was then removed and discarded. ID = internal diameter of the polyvinyl chloride (PVC) tubing. Figure created with BioRender.

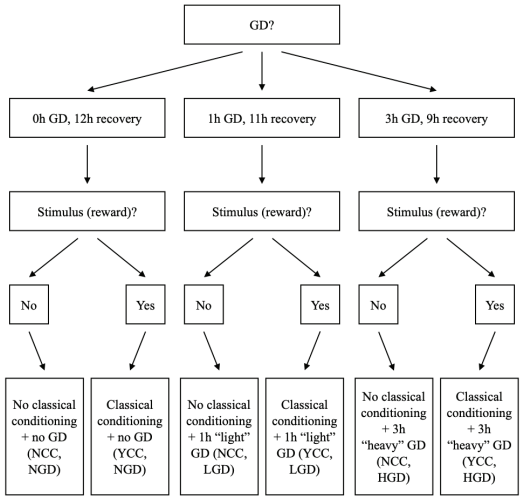


Figure 2: A decision matrix is used to determine experimental group segregation. Each group of fruit flies was created from the above decision matrix, considering whether a stimulus was associated with a reward and whether GD was applied, dividing the sample into 6 experimental groups.

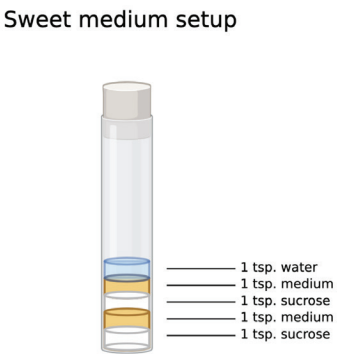


Figure 3: Sweet medium recipe for *D. melanogaster* T-maze classical conditioning training. The sweet medium recipe used as a positive reward was created as shown in the figure, created with BioRender.com.

This setup was used to avoid loss of flies upon transfer between tubes. The aloe vera tube remained a controlled variable, where it was not coated with medium or any form of reward. The apparatus was kept on a 12 h light-dark cycle from 11:00 to 23:00 and remained within temperature ranges of 15–25°C. Due to seasonal and temporal temperature variations, ideal temperature conditions of 21°C could never be firmly established. Flies were kept on medium with an equivalent formula at an identical amount. Additionally, flies were tested at an equivalent age and subjected to matched testing times. Testing occurred across 5 trials in 6 days.

T-maze test:

Initial long-term memory training consisted of anesthetizing fruit flies with CO_2, then transferring flies from their group vials into a tube coated with aloe vera odorant and free of food medium (Figure 4). Flies were given 7 min to acclimate post-transfer: 5 min to recover from anesthesia and 2 min to explore (Figure 5). Once acclimatized, the flies were again anesthetized and transferred into a tube coated with rose hydrosol and containing sweet medium for experimental groups (Figure 5). After the cycle in the rose tube was complete, the flies were anesthetized and transferred into the aloe vera tube, and the process was repeated for a total of 3 rounds (Figure 5). Afterward, flies were anesthetized, transferred back to their group vials, and given 100 min to consolidate memory of the reward with rose odorant, consistent with the classical conditioning method established by Tempel *et al*.[25]

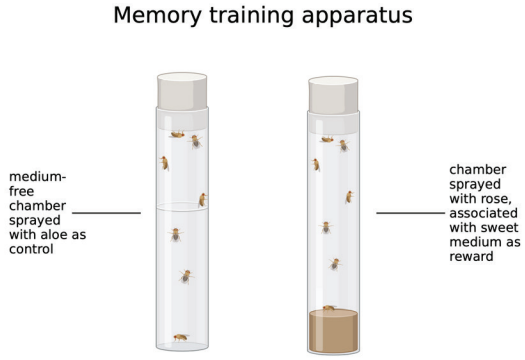


Figure 4: Memory training setup for training *D. melanogaster* for T-maze classical conditioning. Each experimental group of fruit flies was subjected to alternating intervals of 7 min in a medium-free tube coated with aloe odorant or a sweet medium tube coated with rose odorant as shown in the figure, created with BioRender.com.

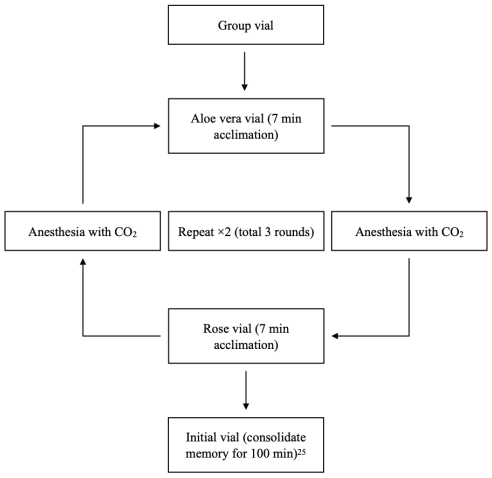


Figure 5: Flow diagram for LTM. Experimental groups of fruit flies had a sweet medium included in the vial with a rose odorant as a reward for classical conditioning. Each LTM training cycle was repeated twice for a total of 3 rounds.

Flies were then immediately subjected to a T-maze test, as shown in Figure 6; the total length of the arms of the T-maze was 10 cm on each arm. The T-maze was then flipped in a minimally biased manner to encourage fruit flies to explore, as they tended to gravitate towards the nearest wall. Each group of flies remained in the T-maze for 5 min to allow navigation towards odorants. Unlike initial testing, where the rose was associated with a reward, this experimental testing did not feature any reward being placed into the rose arm; instead, this relied on classical conditioning to associate the rose with a reward and thereby attract the flies towards the rose odorant.

DOI: 10.36838/IJHSR88.98

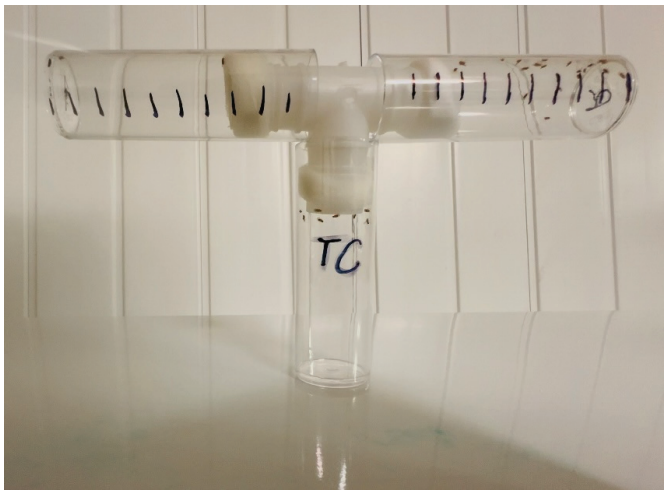


Figure 6: Image of *D. melanogaster* T-maze. Each experimental group of fruit flies was subjected to 5 technical replicates of the T-maze test. Although not clearly seen, A refers to the aloe vera arm, and R refers to the rose arm. TC refers to the transfer chamber, or the tube, in which the fruit flies are initially transferred to within the T-maze.

Immediately afterward, flies in groups that would be subjected to GD were transferred into another chamber free of food medium to deprive glucose and thereby induce ER stress consistent with the findings of García de la Cadena *et al*.[19] GD durations and experimental group stratification are shown in Figure 2. Post-GD, the flies were then immediately transferred into the T-maze containing both odorants for another round of confirmatory memory testing, with 5 min durations per group for testing. Afterward, the flies were euthanized via freezing for 100 h and then disposed of at the local waste plant.

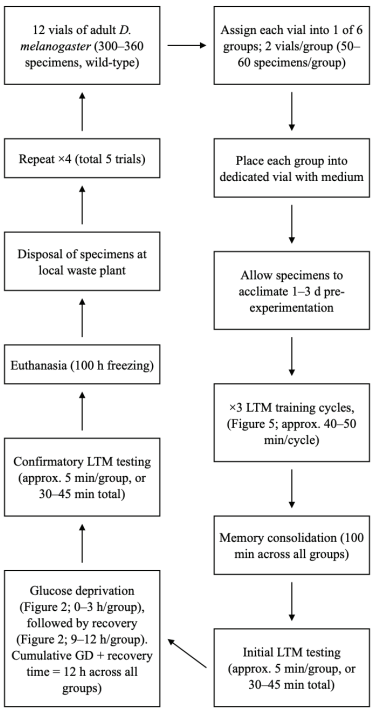


Figure 7: Flow diagram detailing the full experimentation procedure. Each trial followed the generalized timeline from the above flow diagram. Time per trial totaled 970–1015 min from the end of acclimation to the start of euthanasia.

Data analysis:

Digital data was obtained via a stabilized camera recording 4K resolution (3840x2160 pixels) and 60 fps (frames per second) (termed 4K60) video to document the fruit flies' movements. Images were sampled in JPG format at a rate of 1 frame/15 seconds or 4 frames/min via FFmpeg, an open-source video-to-image sampler tool. The frames were then imported into Apple Preview (image/PDF viewer) software, where the distance between the center of the T-maze (baseline) and the fruit fly was manually measured in pixels. The distance fruit flies traveled in horizontal pixels was quantified by converting pixels to cm. Gravitation to a certain odor was quantified by distance in cm from baseline; movement towards the aloe odorant was assigned a negative value, and movement towards the rose odorant was assigned a positive value. Measurements across all flies were averaged to yield a mean value in cm. Flies in the "transfer chamber" (labeled "TC" in Figure 6) were counted as having moved 0 cm as they had not yet explored either odorant-containing arm of the T-maze.

Statistical analyses were performed as follows: one-way ANOVA with Tukey's honestly significant difference (HSD) *post hoc* test applied was used to analyze the variances between experimental groups 1-6, with a p-value α-threshold set to ≤ 0.05. ANOVAs refer to statistical tests analyzing variance between two or more groups and test for significant variances between these groups; Tukey's HSD analysis was used to verify ANOVA results.[26,27] If the *p*-value was below α, there was a significant correlation; if not, there was no significant correlation determined. The independent variable in this experiment was the duration of GD to induce ER stress, and the dependent variable was spatial memory loss quantified in average distance in cm from baseline at the completion of a 5-minute T-maze test (n = 5 technical replicates per experimental group).

■ Results and Discussion

Flies not exposed to classical conditioning towards odorants (NCC, NGD; NCC, LGD; and NCC, HGD) did not navigate further than 0.1 cm from the transfer chamber (TC) on average pre-GD (Figure 8) and 0.3 cm on average post-GD (Figure 9). Odorant-conditioned flies (YCC, NGD; YCC, LGD; YCC, HGD) navigated slightly further, but always within 0.5 cm of TC on average pre- and post-GD (Figure 8; Figure 9). This raw data pre- and post-GD forms the basis for the processed data in Figure 10.

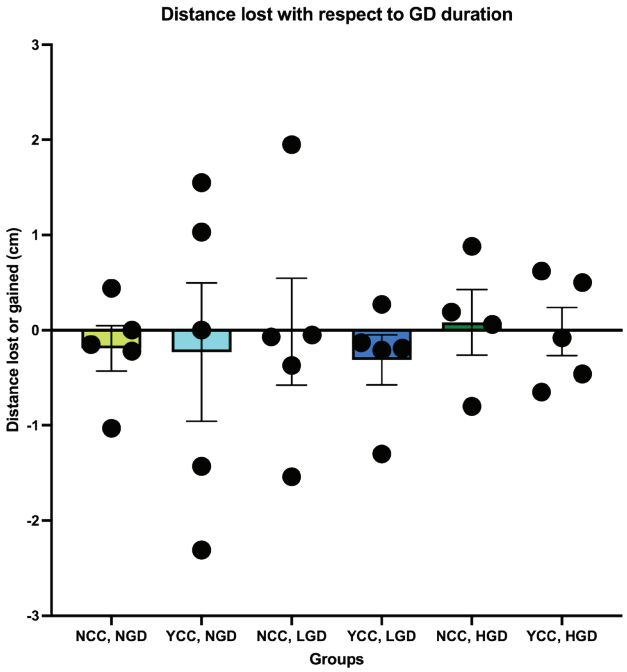


Figure 8: Average movement towards the rose chamber in cm pre-GD. *N* = 5 replicates per group. Negative values denote navigation away from rose or to aloe; positive values denote navigation towards rose. Error bars denote mean ± standard deviation (SD). All groups traveled a varied amount of distance that does not particularly skew towards one group in particular, although groups subjected to classical conditioning generally traveled further in either direction compared to groups not subjected to classical conditioning. Note one outlier in NCC, HGD was invalidated and not counted towards the analysis. Graph created with GraphPad Prism software.

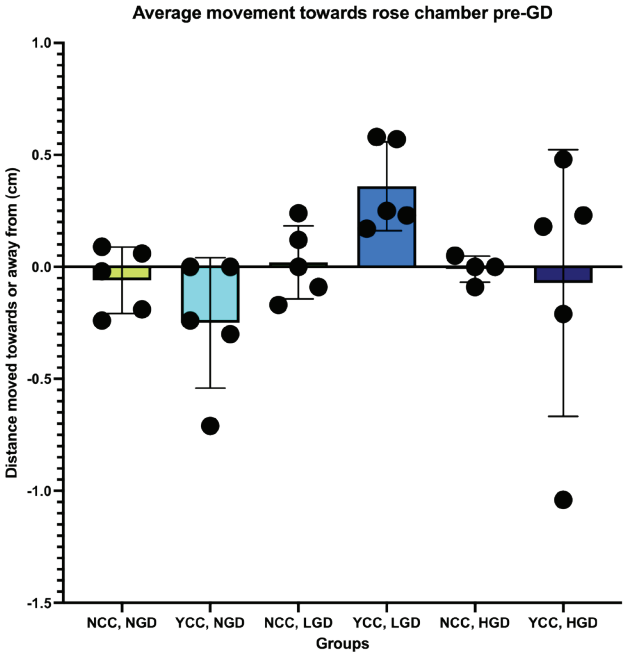


Figure 9: Average movement towards the rose chamber in cm post-GD. *N*= 5 replicates per group. Negative values denote navigation away from rose or to aloe; positive values denote navigation towards rose. Error bars denote mean ± standard deviation (SD). All groups traveled a varied amount of distance that does not particularly skew towards one group in particular. The trend of groups subjected to classical conditioning traveling further in either direction in Figure 8 is not repeated post-GD. Note one outlier in NCC, HGD was invalidated and not counted towards the analysis. Graph created with GraphPad Prism software.

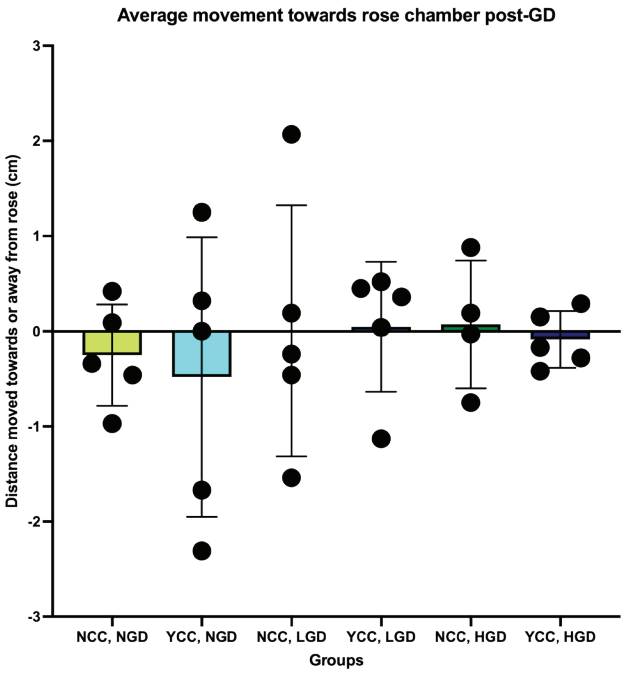


Figure 10: Distance lost with respect to GD duration (post-GD - pre-GD). *N* = 5 replicates per group. Negative values denote navigation away from rose or to aloe; positive values denote navigation towards rose. Error bars denote mean ± standard deviation (SD). All groups lost a varied amount of distance, which does not particularly skew towards one another. An expected trend where YCC, NGD, and YCC, LGD lose a similar amount of distance, whereas YCC, HGD loses a more extreme amount of distance, consistent with the hypothesis, did not materialize. Note one outlier in NCC, HGD was invalidated and not counted towards the analysis. Graph created with GraphPad Prism software.

Flies not exposed to classical conditioning (NCC, NGD; NCC, LGD; NCC, HGD) lost a varied amount of distance that did not exceed –0.2 cm on average; NCC, NGD; NCC, LGD; and NCC, HGD kept consistent with the hypothesis that they did not "lose" memory due to being untrained. Of the experimental groups, YCC and LGD lost the most distance across all 5 trials at an overall loss of –0.31 cm between pre- and post-GD. This was roughly consistent with YCC, NGD, at –0.23 cm between pre- and post-GD. An expected "cliff-shaped" trend where YCC, LGD does not lose significant memory beyond that of YCC, NGD, whereas YCC, HGD loses significant memory, did not materialize; instead, the trend appeared in reverse, where YCC, LGD, and YCC, HGD were comparable, but YCC, HGD lost less spatial memory quantified in distance by average centimeters in the rose arm, at –0.01 cm. Therefore, purely visual documentation appears to refute portions of the initial hypothesis.

In Figure 10, an outlier in NCC, HGD that skewed the values was eliminated; however, it may suggest that fruit flies may be naturally attracted towards roses, particularly after a period of acute starvation, as flies in NCC, HGD experienced 3 h GD but did not undergo classical conditioning. However, it has been shown that the attraction of fruit flies towards decaying fruit or overripe roses in the wild is primarily mediated through their attraction to the scent of yeast (Saccharomyces cerevisiae) fermentation rather than the scent of the fruit or flower.[28]

DOI: 10.36838/IJHSR88.98

The p-value in Tukey's HSD (Table 1), for YCC, NGD vs. YCC, LGD, at >0.9999, was expected, as no significant variance was predicted to be detected owing to the flies being completely non-trained and not directed to gravitate in any direction. However, the *p*-value between groups YCC, NGD vs. YCC, HGD; and YCC, LGD vs. YCC, HGD, at 0.9992 and 0.9964, respectively, did not clear the α-threshold. This appears to refute the second part of my hypothesis that heavy ER stress leads to loss of spatial memory in flies. However, because the hypothesis predicted no significant variance between YCC, NGD, and YCC, LGD, the first component of the hypothesis remains valid.

Heavy (3 h) GD, inducing heavy ER stress, does not induce significant spatial memory loss in fruit flies and does not replicate AD pathogenesis in fruit flies under the selected experimental conditions. Mild ER stress did not induce significant memory loss, as expected.

Table 1: Tukey's honestly significant difference (HSD) on data obtained after one-way ANOVA.

Groups	Mean diff.	95% CI diff.	Below threshold?	Summary	*p**
NCC, NGD vs. YCC, NGD	0.04000	-1.884 to 1.964	No	ns	>0.9999
NCC, NGD vs. NCC, LGD	-0.1760	-2.100 to 1.748	No	ns	0.9997
NCC, NGD vs. YCC, LGD	0.1200	-1.804 to 2.044	No	ns	>0.9999
NCC, NGD vs. NCC, HGD	-0.2745	-2.315 to 1.766	No	ns	0.9982
NCC, NGD vs. YCC, HGD	-0.1780	-2.102 to 1.746	No	ns	0.9997
YCC, NGD vs. NCC, LGD	-0.2160	-2.140 to 1.708	No	ns	0.9992
YCC, NGD vs. YCC, LGD	0.08000	-1.844 to 2.004	No	ns	>0.9999
YCC, NGD vs. NCC, HGD	-0.3145	-2.355 to 1.726	No	ns	0.9965
YCC, NGD vs. YCC, HGD	-0.2180	-2.142 to 1.706	No	ns	0.9992
NCC, LGD vs. YCC, LGD	0.2960	-1.628 to 2.220	No	ns	0.9965
NCC, LGD vs. NCC, HGD	-0.09850	-2.139 to 1.942	No	ns	>0.9999
NCC, LGD vs. YCC, HGD	-0.002000	-1.926 to 1.922	No	ns	>0.9999
YCC, LGD vs. NCC, HGD	-0.3945	-2.435 to 1.646	No	ns	0.9900
YCC, LGD vs. YCC, HGD	-0.2980	-2.222 to 1.626	No	ns	0.9964
NCC, HGD vs. YCC, HGD	0.09650	-1.944 to 2.137	No	ns	>0.9999

Note: Data was fed into GraphPad Prism for ANOVA. Data remains original work. **p* was adjusted.

Experimental limitations:

Although the data remained statistically insignificant, it could be explained by a myriad of limitations. Namely, the T-maze did not seal perfectly, and there was no way to force fruit flies to explore the odorant-containing arms, despite flipping the tube in a minimally biased manner to encourage as many specimens to enter as possible. Therefore, some fruit flies remained stuck in the TC, as opposed to exploring the test arms of the apparatus; these flies were considered to have moved "0" cm per Data Analysis and potentially skewed the data towards no significant changes in distance moved. This phenomenon can be observed in Figure 4.

Central to the insignificance of the data was the provision of 12 h, 11 h, and 9 h recovery time for the fruit flies, where they were inserted into vials with medium (Figure 2), which may have diluted the immediate effects of GD. As the flies were given nutrition immediately post-GD, this constituted a stage where the effects of GD may have been reversed due to ingestion of medium that may have erased its effects in near-entirety. To truly assess the efficacy of GD, a confirmatory memory test should have been performed immediately post-GD as opposed to after a rest period of at least 9 h, which was not considered for this study.

Additionally, fruit flies may have been relatively inactive at the time of testing. Fruit flies have been demonstrated to operate on circadian rhythms akin to those of humans, which may have resulted in fruit flies being unwilling to explore odorants at the time of testing.[29] As initial memory testing occurred during nighttime and confirmatory testing occurred during daytime, it remains possible that any differences in memory could be influenced by their circadian rhythms. Notably, this could explain the lack of change in distance moved in YCC and HGD, as increased activity may have resulted in further willingness to navigate. Furthermore, as fruit flies were starved dietarily, it could further explain the value in YCC, HGD, as flies in that group may have been starved to the extent that they were more willing to navigate to the rose, thinking it brought "food" consistent with the classical conditioning stimulus that the rose equaled sweet medium. Therefore, although GD was the most feasible methodology, other causes of these results, in addition to ER stress, such as metabolic starvation that does not necessarily affect the brain, cannot be ruled out. Any ER stress produced by GD in this study may have affected multiple organ systems, not limited to the brain and central nervous system, which would be more specific for studying AD.

Since the mild (1 h) GD groups (NCC, LGD; YCC, LGD) in this experiment did not experience significant loss of memory beyond that of the untreated groups (NCC, NGD; YCC, NGD), the results could be explained by Fouillet *et al.*,[15] who state a similar result with mild ER stress. This also opens another limitation in that although this study's GD durations technically fell into the bounds of García de la Cadena *et al.*[19] (between 15 min and 4 h of GD), those parameters were only applicable to cell lines and likely did not account for the more sophisticated nature of living organisms. As such, this study likely failed to take into account that living organisms may have needed more GD duration than 3 h to deplete glucose stores. Furthermore, D. melanogaster can produce glucose via gluconeogenesis during periods of nutrient deprivation, and in addition, may have had residual glucose from the training session and resting periods, particularly in the trained groups (groups labeled "YCC"). Notably, neurons may have been able to take advantage of gluconeogenesis, providing a lifeline during GD (Miyamoto & Amrein, 2019).[30] This could have meant that no GD, and consequently ER stress, in the neurons actually occurred, as repositories of glucose remained in neurons even into GD.

At 3 h, the ER stress induced in the "heavy" groups (NCC, HGD; YCC, HGD) may have been mild enough to inhibit and not promote neuronal death, consistent with Fouillet *et al.*[15] Suppose GD were induced for longer, and its induction confirmed molecularly. In that case, the results may resemble those of Lin *et al.*[13] rather than Fouillet *et al.*[15] In addition, mere deprivation of medium remains less reliable than the glucose-free medium method, as indicated by García de la Cadena *et al.*[19] which used a specially formulated glucose-free medium. Purchasing an equivalent in fruit flies would likely have served to resolve some limitations in this study and potentially provide more useful data.

Implications for therapy:

This study may have implications for future research and pathogenic understanding of AD. Namely, this study supports the idea that ER stress is a double-edged sword in AD pathogenesis, in that it may have context-dependent beneficial or detrimental roles in AD development. Studies consistent with this idea have both induced and inhibited ER stress as therapeutic strategies to alleviate symptoms of AD.[10,13,15] It has been shown that post-mortem extirpated human brains from patients with terminal AD displayed increased ER stress in most cases.[10] Care should be taken to ensure that ER stress does not break moderation and become severe or heavy; the definition of which, however, appears to be context-specific.

Most ER stress testing is achieved by tunicamycin or thapsigargin dosage rates per kg of body weight. Still, neither of these drugs is currently authorized for human consumption due to their cellular toxicity.[31,32] As such, future researchers may be served well by implementing an ER stress severity standard that is less cytotoxic. AD medicine may be required to be issued on a case-by-case basis, where doctors analyze the amount of ER stress in the afflicted patient, along with the severity of the disease upon examination.

Untreated diabetes may cause GD because glucose cannot be transported into cells, which is associated with insulin resistance.[22,33] This is the premise of the diabetes-AD hypothesis, or that dementia is T3D, resulting from glucose dyshomeostasis in a brain-specific manner.[19] Examining the linkage between insulin resistance and GD-induced ER stress could therefore be of interest for future researchers, with particular emphasis on modulating glucose administration or deprivation for therapeutic benefit.

Future Directions:

An array of future directions is warranted for researchers aiming to explore the role of ER stress in AD, especially via *in vivo* analyses, in order to generate a more comprehensive understanding of this relatively new field. For one, comparing GD with the effect of biochemical ER stress inducers, especially while adding longer GD durations such as 6 and 12 h, could be worthwhile to see if the induction method for ER stress alters the findings. Additionally, the confirmatory T-maze test should be performed immediately post-subjecting specimens to GD, as by then the immediate effects of GD are more visible as opposed to the 9 h, 11 h, and 12 h rest periods provided for the specimens by this study (Figure 2). Care should also be taken to induce GD-induced ER stress via glucose-free medium and not mere starvation, as the latter is less reliable than the former, as well as possibly confirm the results via biomarkers *in vitro* (for example, relative mRNA and protein expression of XBP-1).

Figure 8 (pre-GD) displays a subtle trend in the direction of the rose odorant in YCC, LGD, although overall findings remain inconclusive. Given this model's relevance in ER stress and glucose dyshomeostasis in AD, it may be beneficial in future studies to include a clear positive control demonstrating robust directional movement. This would help validate the sensitivity and dynamic range of the behavioral assay.

It would also be worthwhile to understand if alleviating ER stress could be used as a complementary treatment strategy in combination with treatments that target the ACH. Researchers should also continue to investigate the pathogenic mechanisms underlying glucose dyshomeostasis and AD symptom onset and progression.

■ Conclusion

Mild and severe ER stress in fruit flies did not significantly impair spatial memory loss following classical conditioning T-maze testing. Although inconclusive, this study may suggest new pathways to studying the role of ER stress as well as GD or glucose dyshomeostasis in AD through ways not previously documented, as well as provide several directions for future scientific and clinical studies to consider. These results indicate a complex, non-linear interaction between ER stress and cognitive performance, supporting the notion that ER stress may influence AD progression in a context-dependent manner. Further experimentation is warranted to establish whether modulation of ER stress could offer therapeutic benefit to AD patients.

■ Acknowledgments

I wish to acknowledge Dr. Anabel Martinez Lyons and Mr. Deighton Zerby for providing me with advice and support as I undertook this study, and my parents for their support as I underwent experimentation.

No generative artificial intelligence (GAI) was used in the formulation of this study or manuscript generation.

■ References

1. Gaugler, J., James, B., Johnson, T., Reimer, J., Scales, K., Tom, S., Weuve, J., & Yeh, J. (2024). 2024 Alzheimer's disease facts and figures. Alzheimer's & dementia: the journal of the Alzheimer's Association, 20(5), 3708–3821. https://doi.org/10.1002/alz.13809
2. United States Census. (2024, November 26). International database. United States Census Bureau. https://www.census.gov/data-tools/demo/idb/#/dashboard?dashboard_page=country&COUNTRY_YR_ANIM=2024
3. Mielke M. M. (2018). Sex and Gender Differences in Alzheimer's Disease Dementia. The Psychiatric Times, 35(11), 14–17.
4. Hardy, J. A., & Higgins, G. A. (1992). Alzheimer's disease: the amyloid cascade hypothesis. Science (New York, N.Y.), 256(5054), 184–185. https://doi.org/10.1126/science.1566067

DOI: 10.36838/IJHSR88.98

5. Levin, J., Vöglein, J., Quiroz, Y. T., Bateman, R. J., Ghisays, V., Lopera, F., McDade, E., Reiman, E., Tariot, P. N., & Morris, J. C. (2022). Testing the amyloid cascade hypothesis: Prevention trials in autosomal dominant Alzheimer's disease. Alzheimer's & dementia: the journal of the Alzheimer's Association, 18(12), 2687–2698. https://doi.org/10.1002/alz.12624
6. Budd Haeberlein, S., Aisen, P. S., Barkhof, F., Chalkias, S., Chen, T., Cohen, S., Dent, G., Hansson, O., Harrison, K., von Hehn, C., Iwatsubo, T., Mallinckrodt, C., Mummery, C. J., Muralidharan, K. K., Nestorov, I., Nisenbaum, L., Rajagovindan, R., Skordos, L., Tian, Y., van Dyck, C. H., … Sandrock, A. (2022). Two Randomized Phase 3 Studies of Aducanumab in Early Alzheimer's Disease. The journal of prevention of Alzheimer's disease, 9(2), 197–210. https://doi.org/10.14283/jpad.2022.30
7. Mintun, M. A., Lo, A. C., Duggan Evans, C., Wessels, A. M., Ardayfio, P. A., Andersen, S. W., Shcherbinin, S., Sparks, J., Sims, J. R., Brys, M., Apostolova, L. G., Salloway, S. P., & Skovronsky, D. M. (2021). Donanemab in early Alzheimer's disease. New England Journal of Medicine, 384(18), 1691–1704. https://doi.org/10.1056/nejmoa2100708
8. Swanson, C. J., Zhang, Y., Dhadda, S., Wang, J., Kaplow, J., Lai, R. Y., Lannfelt, L., Bradley, H., Rabe, M., Koyama, A., Reyderman, L., Berry, D. A., Berry, S., Gordon, R., Kramer, L. D., & Cummings, J. L. (2021). A randomized, double-blind, phase 2B proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-Aβ protofibril antibody. Alzheimer's Research & Therapy, 13(1). https://doi.org/10.1186/s13195-021-00813-8
9. Ajoolabady, A., Lindholm, D., Ren, J., & Pratico, D. (2022). ER stress and UPR in Alzheimer's disease: Mechanisms, pathogenesis, treatments. Cell Death & Disease, 13(8). https://doi.org/10.1038/s41419-022-05153-5
10. Nagar, P., Sharma, P., Dhapola, R., Kumari, S., Medhi, B., & HariKrishnaReddy, D. (2023). Endoplasmic Reticulum stress in alzheimer's disease: Molecular mechanisms and therapeutic prospects. Life Sciences, 330, 121983. https://doi.org/10.1016/j.lfs.2023.121983
11. Chen, X., Shi, C., He, M., Xiong, S., & Xia, X. (2023). Endoplasmic Reticulum Stress: Molecular Mechanism and Therapeutic Targets. Signal Transduction and Targeted Therapy, 8(1). https://doi.org/10.1038/s41392-023-01570-w
12. Salminen, A., Kauppinen, A., Suuronen, T., Kaarniranta, K., & Ojala, J. (2009). Er stress in alzheimer's disease: A novel neuronal trigger for inflammation and alzheimer's pathology. Journal of Neuroinflammation, 6(1), 41. https://doi.org/10.1186/1742-2094-6-41
13. Lin, L., Cao, J., Yang, S. S., Fu, Z. Q., Zeng, P., Chu, J., Ning, L. N., Zhang, T., Shi, Y., Tian, Q., Zhou, X. W., & Wang, J. Z. (2018). Endoplasmic reticulum stress induces spatial memory deficits by activating GSK-3. Journal of cellular and molecular medicine, 22(7), 3489–3502. https://doi.org/10.1111/jcmm.13626
14. Hashimoto, S., Ishii, A., Kamano, N., Watamura, N., Saito, T., Ohshima, T., Yokosuka, M., & Saido, T. C. (2018). Endoplasmic reticulum stress responses in mouse models of Alzheimer's disease: Overexpression paradigm versus knockin paradigm. The Journal of Biological Chemistry, 293(9), 3118–3125. https://doi.org/10.1074/jbc.M117.811315
15. Fouillet, A., Levet, C., Virgone, A., Robin, M., Dourlen, P., Rieusset, J., … Mollereau, B. (2012). ER stress inhibits neuronal death by promoting autophagy. Autophagy, 8(6), 915–926. https://doi.org/10.4161/auto.19716
16. Chow, C. Y., Wolfner, M. F., & Clark, A. G. (2013). Using natural variation in Drosophila to discover previously unknown endoplasmic reticulum stress genes. Proceedings of the National Academy of Sciences of the United States of America, 110(22), 9013–9018. https://doi.org/10.1073/pnas.1307125110
17. Abdullahi, A., Stanojcic, M., Parousis, A., Patsouris, D., & Jeschke, M. G. (2017a). Modeling acute ER stress *in vivo* and *in vitro*. Shock, 47(4), 506–513. https://doi.org/10.1097/shk.0000000000000759
18. ThermoFisher Scientific. (2024, March 30). Safety Data Sheet. ThermoFisher Scientific. https://www.fishersci.com/store/msds?partNumber=AAJ62217MA&productDescription=TUNICAMYCIN,+95%25+10MG&vendorId=VN00024248&countryCode=US&language=en
19. García de la Cadena, S., Hernández-Fonseca, K., Camacho-Arroyo, I., & Massieu, L. (2014). Glucose deprivation induces reticulum stress by the PERK pathway and caspase-7- and calpain-mediated caspase-12 activation. Apoptosis: an international journal on programmed cell death, 19(3), 414–427. https://doi.org/10.1007/s10495-013-0930-7
20. Kalimuthu, K., Kim, J. H., Park, Y. S., Luo, X., Zhang, L., Ku, J. L., Choudry, M. H. A., & Lee, Y. J. (2021). Glucose deprivation-induced endoplasmic reticulum stress response plays a pivotal role in enhancement of TRAIL cytotoxicity. Journal of cellular physiology, 236(9), 6666–6677. https://doi.org/10.1002/jcp.30329
21. Lauretti, E., Li, J. G., Di Meco, A., & Praticò, D. (2017). Glucose deficit triggers tau pathology and synaptic dysfunction in a tauopathy mouse model. Translational psychiatry, 7(1), e1020. https://doi.org/10.1038/tp.2016.296
22. de la Monte, S. M., & Wands, J. R. (2008). Alzheimer's disease is type 3 diabetes-evidence reviewed. Journal of diabetes science and technology, 2(6), 1101–1113. https://doi.org/10.1177/193229680800200619
23. Bronnec, V., & Alexeyev, O. A. (2022). Detailed protocol for germ-free Drosophila melanogaster colonization with Propionibacterium spp. Biofilms. STAR protocols, 3(2), 101342. https://doi.org/10.1016/j.xpro.2022.101342
24. Tempel, B. L., Bonini, N., Dawson, D. R., & Quinn, W. G. (1983). Reward learning in normal and mutant Drosophila. Proceedings of the National Academy of Sciences, 80(5), 1482–1486. https://doi.org/10.1073/pnas.80.5.1482
25. Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2014). Campbell Biology (10th ed.). Pearson Education.
26. Statistics Kingdom. (n.d.). Anova calculator. ANOVA Calculator – One-Way ANOVA and Tukey HSD test. https://www.statskingdom.com/180Anova1way.html
27. Tukey, J. W. (1949). Comparing Individual Means in the Analysis of Variance. Biometrics, 5(2), 99–114. https://doi.org/10.2307/3001913
28. Becher, P. G., Flick, G., Rozpędowska, E., Schmidt, A., Hagman, A., Lebreton, S., Larsson, M. C., Hansson, B. S., Piškur, J., Witzgall, P., & Bengtsson, M. (2012a). Yeast, not fruit volatiles mediate Drosophila melanogaster attraction, oviposition and development. Functional Ecology, 26(4), 822–828. https://doi.org/10.1111/j.1365-2435.2012.02006.x
29. Tataroglu, O., & Emery, P. (2014). Studying circadian rhythms in Drosophila melanogaster. Methods (San Diego, Calif.), 68(1), 140–150. https://doi.org/10.1016/j.ymeth.2014.01.001
30. Miyamoto, T., & Amrein, H. (2019). Neuronal gluconeogenesis regulates systemic glucose homeostasis in Drosophila melanogaster. Current Biology, 29(8). https://doi.org/10.1016/j.cub.2019.02.053
31. Foufelle, F., & Fromenty, B. (2016). Role of endoplasmic reticulum stress in drug-induced toxicity. Pharmacology research & perspectives, 4(1), e00211. https://doi.org/10.1002/prp2.211
32. Jaskulska, A., Janecka, A. E., & Gach-Janczak, K. (2020). Thapsigargin: From Traditional Medicine to Anticancer Drug.

International journal of molecular sciences, 22(1), 4. https://doi.org/10.3390/ijms22010004
33. Nguyen, T. T., Ta, Q. T. H., Nguyen, T. K. O., Nguyen, T. T. D., & Giau, V. V. (2020). Type 3 Diabetes and Its Role Implications in Alzheimer's Disease. International journal of molecular sciences, 21(9), 3165. https://doi.org/10.3390/ijms21093165

Author

Edward Yan is a senior at Westlake High School in Austin, Texas. He is interested in biology, biomedical engineering, and neuroscience. He plans to take the pre-med track in college.

DOI: 10.36838/IJHSR88.98

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

The Prediction of the Movement of Two Initially Stationary Planetary Objects Under Gravity

Juha Shin

Maria Carrillo High School, 6975 Montecito Blvd, Santa Rosa, CA 95409, USA; shin.juhaw0711edu@gmail.com

ABSTRACT: This Paper revisits the classical problem of two point masses initially at rest, moving solely under their mutual Newtonian attraction. By reducing the two-body dynamics to a single relative coordinate in the center-of-mass frame and integrating the energy equation, a closed-form relation between time t and separation R, and the corresponding free-fall time is obtained. Using accepted Earth–Moon parameters, the time to impact from rest is evaluated, and the evolution of the separation R(t), force F(R), and relative acceleration a(R) is plotted. For the practical evaluation of the inverse function R(t), the study constructs an endpoint-constrained polynomial surrogate on a non-dimensional domain and quantifies its uniform error. The surrogate facilitates rapid estimates and serves as an initial guess in numerical solvers. The study also discusses the physical distinction between a point-mass collision at R=0 and a finite-size contact at $R=R_{\oplus}+R_{moon}$, reporting the corresponding time-to-contact. The analysis maintains dimensional consistency and provides all constants, units, and computational details for reproducibility. The results provide a clean test case for educational purposes and for benchmarking numerical integrators in celestial mechanics.

KEYWORDS: Physics, Classical Mechanics, Newtonian Gravitation, Two-Body Problem, Earth–Moon System.

Introduction

The gravitational two-body problem has been central to classical mechanics and astrophysics since Newton's *Principia*.[1] While the general case admits Keplerian orbits, the special case of two initially stationary masses collapsing radially under their mutual attraction remains pedagogically and physically important. It models idealized gravitational free-fall and provides clean benchmarks for validating numerical solvers and understanding dynamical collapse scenarios from planetesimal accretion to binary coalescence.[2–5]

Despite its simplicity, obtaining an explicit time–separation relation for radial infall is nontrivial. For two point masses m_1, m_2 initially separated by R_0 and released from rest, the solution can be written in closed form. Standard derivations exploit the degenerate $e\to1$ limit of Keplerian motion (a radial elliptic trajectory). The relevant formulas are available in the literature,[6–8] but here these results are re-derived for completeness and clarity. In particular, the study first derives the exact expressions for two-body radial free-fall (time–distance relation and free-fall time), applies them to the Earth–Moon system using accepted parameters (gravitational constant G, planetary masses, and mean Earth–Moon separation), and constructs a smooth polynomial surrogate for R(t) constrained at the endpoints and optimized in a least-squares sense, enabling fast evaluation while retaining uniform accuracy across the interval. Beyond serving as a computational proxy, the surrogate clarifies sensitivity to initial separation and total mass, and it is readily embedded as an initial guess in root-finding schemes for the inverse problem.

The remainder of this paper is organized as follows. Section 1 (Methods) presents the equations of motion and the exact time–distance relation for free-fall. Section 2 (Results and Discussion) evaluates the Earth–Moon free-fall from rest and compares exact vs. polynomial results using standard constants. Section 3 discusses accuracy, physical implications, and limitations (finite sizes, tidal effects, non-point-mass deviations). Section 4 concludes the paper and outlines extensions to nonzero angular momentum and perturbed *N*-body settings.

Methods

Considering two-point masses initially at rest and separated by a distance R_0 , Newtonian gravity (inverse-square law) predicts they would accelerate toward each other. Below, the equations of motion are presented, the time-distance relation is verified via energy integration, and finite-size effects and a dimensionless check are included. All derivations are made fully explicit.

Equations of motion and reduction to relative coordinates:

Consider two point masses m_1 and m_2 located at positions r_1 and r_2, separated by $R=|r_2-r_1|$. Newton's law of gravitation provides the force on each mass, $F_{12}=-F_{21}=-\frac{Gm_1m_2}{R^2}\hat{R}$, where $\hat{R}$ is the unit vector from m_1 to m_2. Newton's second law then implies $m_1\ddot{r}_1=F_{12}$, $m_2\ddot{r}_2=F_{21}$ Subtracting these equations and defining the relative coordinate $R=r_2-r_1$ yields $\ddot{R}=-\frac{G(m_1+m_2)}{R^2}\hat{R}$. In the center-of-mass frame, the dynamics reduce to the evolution of the scalar separation R(t). Newton's second law, together with the mutual gravitational force, therefore gives

$$\ddot{R} = -\frac{G(m_1 + m_2)}{R^2}, \qquad (1)$$

which is the radial equation of motion governing the relative infall of the two masses. The force acts inward along the line joining the masses. A dot denotes a time derivative, and the reduced mass formulation is used to obtain an effective one-body

DOI: 10.36838/IJHSR88.107

equation with total mass m_1+m_2. Equation (1) is equivalent to the motion of a unit-mass particle moving radially in the gravitational potential $-G(m_1+m_2)/R$, with relative acceleration $a(R)=\ddot{R}$ given by the right-hand side of Equation (1).

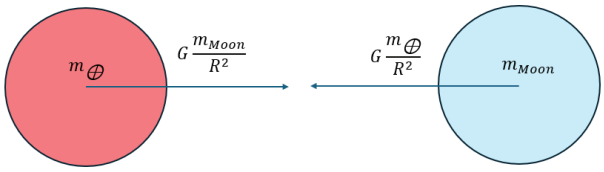


Figure 1: Schematic of the initial state with Earth ($m_{\oplus}$, left) and Moon (mmoon, right) at rest, separated by R_0. Arrows indicate the directions of initial accelerations.

Illustrated in Figure 1 is a schematic of the initial configuration for the Earth–Moon case: the Earth of mass $m_{\oplus}$ (red sphere, left) and the Moon of mass mmoon (blue sphere, right) start at rest, separated by R_0. Arrows indicate the directions of the initial accelerations. In analysis, both bodies are treated as point masses; a collision occurs when $R\to 0$ (or when physical contact is made at finite radii $R_{\oplus}+R_{moon}$ in a finite-size model). Each body's initial acceleration magnitude is Gm_{other}/R_0^2 , where m_{other} is the mass of the companion body in the two-body system (Earth's arrow is shorter in the figure to qualitatively reflect its smaller acceleration due to $m_{\oplus}\gg m_{moon}$).

Energy integral and closed-form solution:

Multiplying both sides of Equation (1) by $\dot{R}$ gives $\dot{R}\ddot{R}=-\frac{G(m_1+m_2)}{R^2}\dot{R}$. The left-hand side can be written as a total derivative, $\dot{R}\ddot{R}=\frac{d}{dt}\left(\frac{1}{2}\dot{R}^2\right)$, while the right-hand side satisfies $-\frac{G(m_1+m_2)}{R^2}\dot{R}=\frac{d}{dt}\left(-\frac{G(m_1+m_2)}{R}\right)$. Integrating both sides with respect to time, therefore, gives the conservation of mechanical energy: $\frac{1}{2}\dot{R}^2-\frac{G(m_1+m_2)}{R}=\text{constant}$. Imposing the initial condition $\dot{R}(0)=0$ at $R(0)=R_0$ fixes the integration constant to be $-\frac{G(m_1+m_2)}{R_0}$. Solving for $\dot{R}$ then gives

$$\dot{R}^2 = 2G(m_1+m_2)\left(\frac{1}{R}-\frac{1}{R_0}\right). \qquad (2)$$

which is the closed-form expression for radial speed as a function of separation. Here, the negative root (not shown) would correspond to R increasing in time; the minus sign is chosen for $\dot{R}$ to enforce decreasing R(t) as the masses move toward each other. Equation (2) can be rearranged and separated as $dt=-dR/\sqrt{2G(m_1+m_2)(1/R-1/R_0)}$. Integrating from R_0 (at t=0) to a general separation R (at time t) gives the implicit time–distance relation. Evaluating the integral, one obtains a closed-form solution:[6–8]

$$t(R) = \sqrt{\frac{R_0^3}{2G(m_1+m_2)}}\left[\arccos\sqrt{\frac{R}{R_0}}+\sqrt{\frac{R}{R_0}\left(1-\frac{R}{R_0}\right)}\right]. \qquad (3)$$

Setting R=0 in Equation (3) yields the collision time (free-fall time from R_0 to 0):

$$t_{ff} \equiv t(R=0) = \frac{\pi}{2\sqrt{2}}\sqrt{\frac{R_0^3}{G(m_1+m_2)}}. \qquad (4)$$

For example, substituting $R_0=3.844\times10^8$ m (mean Earth–Moon distance) and $m_1+m_2=m_{\oplus}+m_{moon}$ into Equation (4), one finds $t_{ff}\approx1.1576\times10^5$ s≈115.76 hours≈4.82 days (assuming point masses).[8] If instead the bodies have finite radii $R_{\oplus}$ and R_{moon}, physical contact occurs at $R=R_{\oplus}+R_{moon}$ before the point-mass collision. The corresponding time-to-contact is obtained by evaluating Equation (3) at $R=R_{\oplus}+R_{moon}$:

$$t_{contact} = \sqrt{\frac{R_0^3}{2G(m_1+m_2)}}\left[\arccos\sqrt{\frac{R_{\oplus}+R_{moon}}{R_0}}+\sqrt{\frac{R_{\oplus}+R_{moon}}{R_0}\left(1-\frac{R_{\oplus}+R_{moon}}{R_0}\right)}\right].$$

This is slightly less than t_{ff} because the bodies do not need to fall all the way to R=0. In the Earth–Moon case, using $R_{\oplus}\approx6.371\times10^6$ m and $R_{moon}\approx1.737\times10^6$ m , one finds $t_{contact}\approx1.1544\times10^5$ s≈115.40 hours (about 0.36 hours shorter than t_{ff}). Reporting both t_{ff} and $t_{contact}$ clarifies the difference between an ideal point-mass collision and a physical impact in finite-size bodies. Notably, the inverse-square singularity at R=0 is avoided in the finite-size model since the masses never get closer than $R_{\oplus}+R_{moon}$.

Equation (3) is a monotonic decreasing function of R on $(0,R_0]$ (with $t\to0$ as $R\to R_0$ and $t\to t_{ff}$ as $R\to0$). This monotonicity means the relation can be inverted to express R as a function of t. In general, the inverse R(t) cannot be written in elementary closed form, but it can be obtained by numerical root-finding. In the next section, the study constructs an approximate analytic surrogate for R(t) to avoid iterative inversion.

Non-dimensionalization and early-time check:

It is convenient to introduce dimensionless variables as a sanity check for limiting behavior. Let $\tau=t/t_{ff}$ be the normalized time (with $\tau=1$ at collision) and $y=R/R_0$, the normalized separation (with $y(0)=1$ and $y(1)=0$). Equation (3) can be rewritten in dimensionless form as

$$\tau(y) = \frac{2}{\pi}\left[\arccos\sqrt{y}+\sqrt{y(1-y)}\right], \qquad (5)$$

which defines the normalized time-separation relation.

For early times ($\tau\ll1$), setting $y=1-\varepsilon$ with $\varepsilon\ll1$ gives $y(\tau)\approx1-\frac{\pi^2}{16}\tau^2$ obtained from the small-ε expansions of $\arccos(\sqrt{y})$ and $\sqrt{y(1-y)}$. In dimensional terms, $R(t)\approx R_0\left(1-\frac{\pi^2}{16}\frac{t^2}{t_{ff}^2}\right)$ implies an initial acceleration $a(0)=G(m_1+m_2)/R_0^2$, matching the direct Newtonian acceleration from rest and confirming consistency at $t\to0$. one can expand $y(\tau)$ to find $y\approx1-1/2\tau^2$ (using $\arccos\sqrt{y}\approx\pi/2-\sqrt{y}$ for y near 1). In dimensional terms, $R(t)\approx R_0(1-1/2\omega^2t^2)$ for small t, which implies initial acceleration $a(0)=\omega^2R_0=G(m_1+m_2)/R_0^2$. This matches the direct Newtonian acceleration from rest, confirming consistency at $t\to0$. The variables used in this study are summarized in Table 1.

Table 1: Variable used in the Non-dimensionalization and early-time check.

Symbol	Definition	Meaning
τ	t/t_{ff}	Normalized time (0→1)
y	R/R_0	Normalized separation (1→0)
ω	$\frac{\pi}{2\sqrt{2}t_{ff}}$	Constant

Numerical inversion and surrogate polynomial:

For arbitrary times $0<t<t_{ff}$, one may invert Equation (3) numerically to obtain R(t). A robust approach is to solve $t(R)=t_{given}$ by bracketing and root-finding (e.g., using Brent's

method) at each desired t. In practice, since t(R) is monotonic, this inversion is straightforward. Tolerances (absolute and relative error) can be set to ensure the solution R(t) is obtained to high precision; in our computations, double-precision arithmetic is used, and iterations are terminated when relative changes fall below 10^{-12}. Figure 2 shows the resulting R(t) for the Earth–Moon example, obtained by numerically inverting the exact implicit relation.

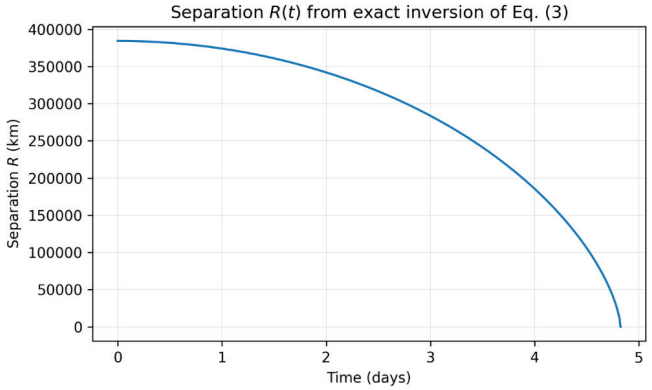


Figure 2: Separation R(t) obtained by numerically inverting the exact time–distance expression, for two bodies initially at rest (Earth–Moon parameters). The separation starts at $R_0 \approx 3.844\times10^5$ km and decreases to zero over $t_{ff} \approx 4.82$ days. For reference, the early-time behavior follows $y \approx 1-(\pi^2/16)\tau^2$, consistent with the Newtonian acceleration from rest.

Although numerical inversion is efficient, it is sometimes helpful to have an analytic approximation for R(t). A polynomial surrogate $R_{approx}(t)$ is constructed on $0 \le t \le t_{ff}$ that satisfies the known boundary values $R(0)=R_0$ and $R(t_{ff})=0$ and minimizes the error in between. To avoid Runge oscillations, a Chebyshev polynomial basis is often recommended for approximating smooth functions on a finite interval;[15,16] however, for simplicity, a monomial basis with endpoint constraints is employed in this work. A least-squares fit of the form $y(\tau)=\sum_{n=0}^{N} a_n \tau^n$ (with $y(0)=1$ and $y(1)=0$ enforced) is performed over a dense grid of $\tau=t/t_{ff} \in [0,1]$. The polynomial degree N can be increased until the approximation is uniformly accurate to the desired tolerance.

In our Earth–Moon case, it is found that N=14 (degree 14) is sufficient to achieve better than ~1.4% accuracy in R throughout the fall. Figure 3 shows the approximate inversion $R_{approx}(t)$ (dashed curve) versus the exact solution R(t) (solid curve); the maximum absolute error is about 1.12% of R_0. The surrogate polynomial coefficients $a_0,\ldots,a_{14}$ are listed in Table 4 (Appendix). Such a surrogate allows rapid evaluation of R(t) without iterative solvers and can serve as an initial guess if one needs to solve for t(R) via root-finding.

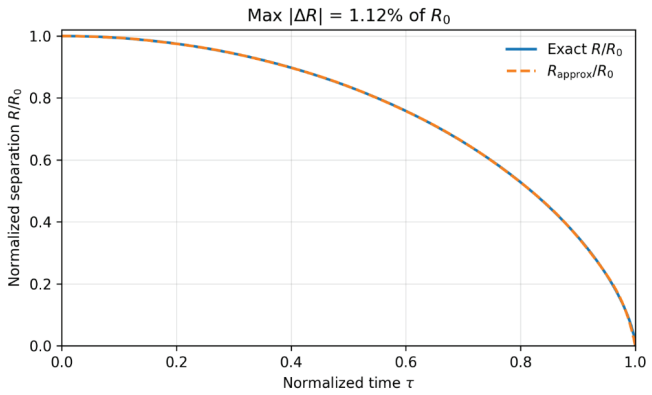


Figure 3: Approximate inverted solution $R_{approx}(t)$ (dashed line) versus the exact solution R(t) (solid line) for a two-body free-fall from rest. The polynomial surrogate (degree 14) is constructed on $0 \le t \le t_{ff}$ and yields a maximum error of ~1.12% in R. The curves are nearly indistinguishable at this scale. The inset text reports the maximum absolute error as a percentage of R_0.

Results and Discussion

Earth–Moon free-fall example:

Using the formulas from Section 1, the free-fall time and trajectory can be computed for specific masses. The physical parameters used are $G=6.67430\times10^{-11}$ $m^3\cdot kg^{-1}\cdot s^{-2}$ (CODATA 2018),[11] $m_\oplus=5.972\times10^{24}$ kg, $m_{moon}=7.342\times10^{22}$ kg (so $m_1+m_2 \approx 6.045\times10^{24}$ kg),12 and $R_0=3.844\times10^8$ m (the average Earth–Moon center distance).[13] For the Earth–Moon system, the physical parameters used in this study are summarized in Table 2.

Table 2: Physical parameters used in the Earth-Moon free-fall calculation.

Quantity	Symbol	Value	Units	Source
Gravitational constant	G	6.67430×10^{-11}	$m^3\cdot kg^{-1}\cdot s^{-2}$	CODATA 2018
Earth mass	$m_\oplus$	5.972×10^{24}	kg	JPL GM values
Moon mass	m_{moon}	7.342×10^{22}	kg	JPL GM values
Initial separation	R_0	3.844×10^8	m	NASA Fact Sheet

Substituting these values into Equation (4) confirms a free-fall time $t_{ff} \approx 4.82$ days, as noted above. Figure 2 plots the separation R(t) over this interval. Starting from $R_0 \approx 3.844\times10^5$ km at t=0, the separation decreases slowly at first (due to weak initial gravity) and then more rapidly, approaching zero in $t_{ff} \approx 4.82$ days. Figure 4 shows the inverse relation t(R): the time required to reach a given separation R. For example, from R_0 down to 0.5 R_0 (half the initial distance), the elapsed time is about 3.62 days, whereas the remaining fall from 0.5 R_0 to 0 takes only about 1.2 days, reflecting the stronger gravitational acceleration at smaller separations. The initial half of the fall thus takes significantly longer, consistent with the gradual early decrease in R.

DOI: 10.36838/IJHSR88.107

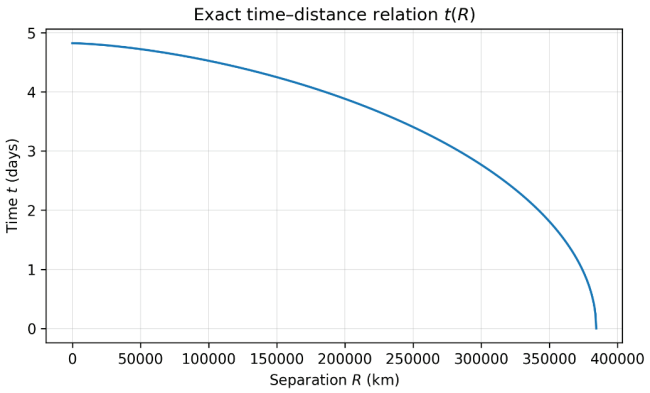


Figure 4: Exact time–distance relation t(R) for radial infall from rest (Earth–Moon parameters). This is the functional inverse of Figure 2. Near the initial separation (R close to R_0), time increases slowly, but as R becomes small, t(R) grows steeply (diverging as R→0 for point masses). For a finite contact radius ($R_\oplus+R_{moon}$), the divergence is avoided, and the contact occurs at a finite time (≈4.81 days in this case).

Newton's law gives the gravitational force between the masses $F=Gm_1m_2/R^2$. Starting from an initial value $F_0=Gm_1m_2/R_0^2$, the force grows without bound as R→0 in the point-mass model. Figure 5 plots F as a function of separation R. On a log–log scale, it appears as a straight line, since $F\propto R^{-2}$ (slope -2 in the figure). For the Earth–Moon case, the initial force is $F_0\approx1.99\times10^{20}$ N; by the time the separation has shrunk to 10^7 m, the force exceeds 10^{23} N. However, in reality, physical contact would occur before such extreme forces are reached. In our finite-size model, the force at contact ($R=R_\oplus+R_{moon}$) is $F_{contact}\approx4.65\times10^{22}$ N, which is immense but finite. After contact, the point-mass idealization breaks down, and other effects (material deformation, fragmentation, etc.) come into play, beyond the scope of this work.

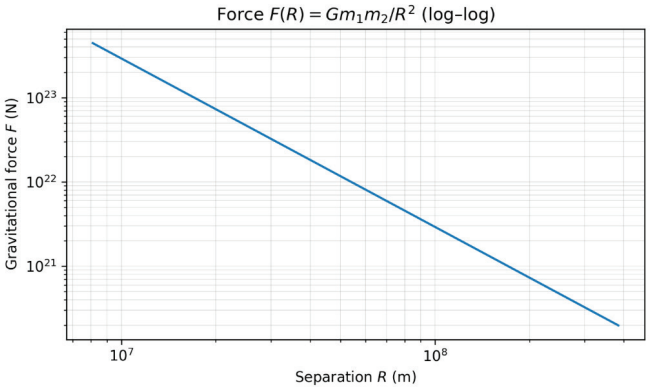


Figure 5: Gravitational force $F=Gm_1m_2/R^2$ versus separation R for the two-body infall (log–log scale). The force increases as R^{-2} (straight line of slope –2 on the log–log plot). For the Earth–Moon parameters, F starts at 2×10^{20} N and grows to 10^{23} N as R approaches 10^7 m. The open circle indicates the force at Earth–Moon contact ($R=R_\oplus+R_{moon}$); the divergence for R→0 is avoided in a finite-size model.

The relative acceleration of the two bodies is $a(R)=\ddot{R}=G(m_1+m_2)/R^2$ (directed inward). This equals F/μ, where $\mu=m_1m_2/(m_1+m_2)$ is the reduced mass. As R decreases, a(R) increases to R^{-2}, similar to F. Figure 6 shows a(R) vs. R on a log–log plot. At the initial distance R_0, the acceleration is $a(0)=G(m_1+m_2)/R_0^2$. For the Earth–Moon system, this is $a(0)\approx2.73\times10^{-3}$ m/s² (about 2.8×10^{-4} g). By the time R reaches 10^7 m, the relative acceleration has grown to 6 m/s² (over half of Earth's surface gravity). At contact ($R_\oplus+R_{moon}$), the acceleration is 9.81 m/s², coincidentally, about 1 g. In the point-mass model a(R) would diverge as R→0, but this unphysical divergence is resolved by the finite-size cutoff.

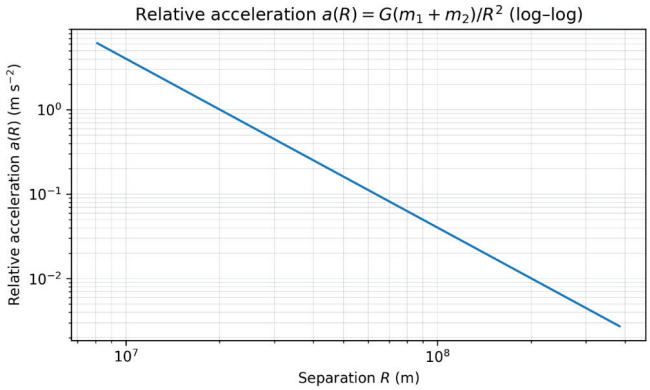


Figure 6: Relative acceleration $a(R)=G(m_1+m_2)/R^2$ versus separation (log–log scale). Like the force in Figure 5, the acceleration follows an R^{-2} power law. For reference, Earth's surface gravity (9.81 m/s²) is indicated by the horizontal line; the relative acceleration exceeds 1 g once the two bodies are within 8×10^6 m of each other.

Representative force and acceleration value at selected separations for the Earth-Moon free-fall calculation are summarized in Table 3.

Table 3: Physical parameters used in the Earth-Moon free-fall calculation.

Separation R	Force F(R)[N]	Acceleration a(R)[m/s²]	Notes
$R_0=3.844\times10^8$m	$\approx2.0\times10^{20}$	$\approx2.73\times10^{-3}$	Initial state
1.0×10^7m	$\approx1.0\times10^{23}$	≈6	Mid-infall
$R_\oplus+R_{moon}$	$\approx4.65\times10^{22}$	≈9.81	Contact

Accuracy of the surrogate polynomial:

Figure 3 already demonstrated that the degree-14 polynomial $R_{approx}(t)$ is almost indistinguishable from the exact inversion of R(t) at the scale of the full trajectory. The maximum discrepancy (at τ≈0.8) is about 1.12% of R_0. The error function $\Delta R(t)=R_{approx}(t)-R(t)$ can then be examined in more detail; it starts at 0 at t=0 (by construction), remains under 5×10^{-3} R_0 for most of the interval, then rises to $\approx0.0112R_0$ around t≈0.8 tff, before returning to 0 at $t=t_{ff}$ (again by construction). The signed error is positive in the middle portion (meaning R_{approx} slightly overestimates R during that time). By integrating or averaging the error, one can confirm that the surrogate is unbiased (the positive and negative deviations cancel out, as expected from a least-squares fit).

The surrogate polynomial, although not exact, is useful for rapid calculations or as an analytic proxy. For instance, one can differentiate $R_{approx}(t)$ to get an approximate v(t) or a(t) without dealing with square roots or trigonometric functions. Moreover, if one needed t as a function of R, plugging $R_{approx}(t)$ into Equation (3) and iterating would converge faster than starting from a trivial initial guess. In summary, the surrogate captures the main behavior of the exact solution with modest error

(<1.5%) and illustrates how the dynamics depend on parameters in a smooth, closed-form way.

Physical implications and limitations:

The derived expressions assume point masses moving under mutual gravity alone. However, several factors, such as tidal forces, forces from external bodies, material strength, and energy dissipation, are neglected for simplicity, which deform or fragment physical bodies long before R=0. For instance, Earth and Moon would experience tidal disruption and heating well before collision. Our finite-size contact time $t_{contact}$ (evaluated at $R_{\oplus}+R_{moon}$) provides a more physical estimate of the infall duration than tff; in the Earth–Moon case, the difference is only 0.3 hours (0.3% of the total time), but for less massive or more diffuse bodies, the discrepancy could be larger. In addition, in the Earth–Moon scenario, the presence of the Sun, other planets, or any initial motion (angular momentum) is ignored. Thus, our results apply to an idealized, isolated two-body system. For the actual Earth–Moon system, which formed from a disk after a giant impact,[14] such free-falls never occurred. However, one can imagine creating two stationary masses in deep space and letting them go; our formulas would describe their subsequent motion precisely, up until the moment of contact.

Finally, general relativity is neglected. With extreme accelerations and small separations in the late stages of infall, relativistic corrections (periastron shift, gravitational radiation) could play a role, especially for very massive bodies (e.g., black holes). Our classical treatment would break down in those regimes. However, for most practical cases like planetesimals or stars on quasi-radial trajectories, Newtonian gravity provides an excellent approximation.[6,7]

In summary, the point-mass free-fall problem provides a baseline scenario that is solvable in closed form. Deviations due to non-zero size, tidal forces, additional bodies, or relativistic effects can be considered as perturbations to this baseline. The clarity of the solution makes it a useful test case for code verification or classroom demonstration. Any N-body integrator, for instance, should reproduce the R(t) curve of Figure 2 when applied to two bodies from rest (to within numerical error).

■ Conclusion

This paper presents the derivation of the radial free-fall problem for two initially stationary masses and obtains the corresponding time–distance relation and dynamics. For Earth–Moon parameters, the free-fall (collision) time from the initial separation is $t_{ff}\approx115.76$ hours (≈4.82 days), and the corresponding distance–time and force–distance profiles were plotted (Figures 3–6). A smooth degree-14 polynomial inverse serves as a practical surrogate $R_{approx}(t)$, achieving about ~1.12% maximum error across the entire interval, demonstrating that an analytic fit can accurately approximate the implicit solution.

This study, while idealized, offers insight into the nature of gravitational infall and can serve as a benchmark. The formulas derived can be used to check the correctness of numerical integrators (e.g., one can compare a simulation's output to Equation (3) or Equation (4)). Future work could extend these results to cases with nonzero angular momentum, leading to orbital motion rather than direct collision, or to few-body settings where no closed-form solution exists, but similar numerical or surrogate approaches may be applied. The impact of additional forces, such as drag or radiation pressure, could also be considered to see how the free-fall time deviates from the simple estimates given here.

■ Acknowledgments

The author thanks James (Brown University) for guiding the overall flow of this manuscript and for many helpful discussions. Appreciation is extended to the teachers, counselors, and classmates for their encouragement, whose questions often shaped the way the author approached the derivations. The author is additionally grateful to Maria Carrillo High School (Santa Rosa, CA) for library and computing resources used in this work. Any remaining errors are the author's responsibility.

■ References

1. Newton, I. Philosophiæ Naturalis Principia Mathematica; Royal Society: London, 1687.
2. Goldstein, H.; Poole, C.; Safko, J. Classical Mechanics, 3rd ed.; Addison-Wesley: San Francisco, 2002.
3. Landau, L. D.; Lifshitz, E. M. Mechanics, 3rd ed.; Pergamon: Oxford, 1976.
4. Roy, A. E. Orbital Motion, 4th ed.; Institute of Physics Publishing: Bristol, 2005.
5. Battin, R. H. An Introduction to the Mathematics and Methods of Astrodynamics, Rev. ed.; AIAA: Reston, VA, 1999.
6. Foong, S. K. From Moon-fall to motions under inverse square laws. Eur. J. Phys. 2008, 29, 987–1003.
7. Mungan, C. E. Radial motion of two mutually attracting particles. Phys. Teach. 2009, 47, 502–507.
8. Free-fall time; inverse-square radial infall. Wikipedia. https://en.wikipedia.org/wiki/Free-fall_time (accessed Sept 28, 2025).
9. Vallado, D. A. Fundamentals of Astrodynamics and Applications, 4th ed.; Microcosm Press/Springer: Hawthorne, CA, 2013.
10. Bate, R. R.; Mueller, D. D.; White, J. E. Fundamentals of Astrodynamics; Dover: New York, 1971.
11. Mohr, P. J.; Taylor, B. N.; Newell, D. B. CODATA recommended values of the fundamental physical constants: 2018. Rev. Mod. Phys. 2018, 88, 035009.
12. Jet Propulsion Laboratory. Astrodynamic Parameters (GM values): Earth GM ≈3.986004355×10^5 $km^3{\cdot}s^{-2}$, Moon GM ≈4902.8001 $km^3{\cdot}s^{-2}$. https://ssd.jpl.nasa.gov/astro_par.html (accessed Sept 28, 2025).
13. NASA Science. Moon — Moon Fact Sheet: Average Earth–Moon Distance 384,400 km. https://nssdc.gsfc.nasa.gov/planetary/factsheet/moonfact.html (accessed Sept 28, 2025).
14. Park, R. S.; Folkner, W. M.; Williams, J. G.; Boggs, D. H. The JPL Planetary and Lunar Ephemerides DE440 and DE441. Astron. J. 2021, 161, 105.
15. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, 2007. (Chapter on Chebyshev approximation.)
16. Trefethen, L. N. *Approximation Theory and Approximation Practice*; SIAM: Philadelphia, 2013.

DOI: 10.36838/IJHSR88.107

Authors

Juha Shin is a student researcher at Maria Carrillo High School (Santa Rosa, CA). His work focuses on Newtonian gravitation, two-body dynamics, and polynomial approximations for inverse problems. He plans to extend this study to few-body simulations and to incorporate non-gravitational effects in future research.

Appendix

Polynomial coefficients for $R_{approx}(t)$:

The fitted coefficients for the degree-14 surrogate polynomial $y(\tau)=R_{approx}(t)/R_0=\Sigma_{n=0}^{14} a_n \tau^n$ (least-squares fit over $0\le\tau\le1$ with $y(0)=1$, $y(1)=0$) are given in the table below:

Table 4: Polynomial coefficients for $R_{approx}(t)$.

Coefficient	Value (Earth–Moon fit)
a_0	9.9923×10^{-1}
a_1	1.8753×10^{-1}
a_2	-1.1317×10^{1}
a_3	2.6174×10^{2}
a_4	-3.4673×10^{3}
a_5	2.8060×10^{4}
a_6	-1.4883×10^{5}
a_7	5.3969×10^{5}
a_8	-1.3699×10^{6}
a_9	2.4570×10^{6}
a_{10}	-3.0980×10^{6}
a_{11}	2.6858×10^{6}
a_{12}	-1.5238×10^{6}
a_{13}	5.0936×10^{5}
a_{14}	-7.6033×10^{4}

These coefficients correspond to the polynomial depicted in Figure 3. The uniform error of this fit is approximately 1.12% of R_0, occurring around $\tau\approx0.8$.

IJHSR International Journal of High School Research

is a publication of

terra science and education

N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education