

Nikola Tesla STEM High School

Redmond, Washington



2023-2024 Research Journal

Introduction

This is the ninth edition of the Nikola Tesla STEM High School Research Journal which enables our students to share their research with the academic and business communities. The investigation included in this edition demonstrates both the increasing depth and diversity of our studies. In addition, we have included a list of the awards and honors earned by Nikola Tesla STEM students this academic year.

We express our sincere gratitude to the dedicated faculty and staff for ensuring that all students are afforded every opportunity to showcase their knowledge and talents and for their support in the publication of this journal. We are especially grateful to Kate Allender, the Research Paper Review Board, our student researchers and their mentors for their dedication to STEM Literacy.

Journal Staff

Editor in Chief: Ahanya John

Managing Editor: Grace Zhao

Design Editor: Sanskriti Sinha

Section Editors, Research Papers: Aarush Kapoor, Shreeyaa Muhunthan, Keerthi Rajesh, Shailja Tyagi

Section Editors, Awards: Pranag Ambekar, Akchara Mukunthu

Journal Support: Kate Allender, Amie Karkainen, Debi Welcome

Current and past Research Journals are available online at: <https://tesla.lwsd.org/academics>

Nikola Tesla STEM High School

Lake Washington School District

4301 228th Ave. NE, Redmond, Washington 98053

(425) 936-2770 * <https://tesla.lwsd.org>

Principal: Amie Karkainen

Associate Principal: Sumeyye Cardakli

TABLE OF CONTENTS

BEHAVIORAL SOCIAL SCIENCES	1
<i>Color Me Surprised Evaluating and Identifying the Impacts of Logo Color Trends on Business Brand Awareness</i>	2
<i>Competitive Learning & Cognitive Resilience: Analyzing Peers' Influence on Planning and Problem Solving Skills</i>	13
<i>Eliminating Cross-Race Bias in Witness Scenarios through Childhood Exposure</i>	23
<i>Interplay of Individual Personality and Language Background on Speech Accommodation</i>	32
<i>Is it Actually That Dang Phone? The Effect of Social Media Usage on Emotional Response to Art</i>	39
<i>Optimizing 2000m Rowing Performance: An Experimental Study Assessing Ergometer Performance in Relation to Pacing Strategy</i>	58
<i>Testing the effectiveness of progressive muscle relaxation on the cognitive performances of adolescents with somatic anxiety</i>	65
BIOMEDICAL, CELLULAR BIOLOGY	75
<i>Autotoxicity of Essential Oils on Photosynthesis in Mentha spicata and Thymus vulgaris</i>	76
<i>Investigating the Success of Microbial Treatment for Gluten Digestion</i>	83
<i>Supplementing Immunity: Investigating the Effect of Vitamin D Supplementation on Human Cancer T-cell Proliferation and IL-2 Secretion</i>	92
<i>Surviving Salt: The Osmoprotective Effects of Exogenous Trehalose and Sucrose on Mung Bean (Vigna radiata) Growth Under NaCl Stress</i>	102
<i>The Antibacterial Effect of Ohi'a Lehua Blossom Honey on Escherichia coli and Bacillus cereus</i>	110
COMPUTATIONAL BIOLOGY, ARTIFICIAL INTELLIGENCE	122
<i>A Novel Approach for Detection and Quantification of Stool on Abdominal Radiographs in Children using Machine Learning</i>	123
<i>Exploring the Influence of AI-Generated Imagery on E-Commerce Consumer Behavior</i>	129
<i>Implementation of a Convolutional Neural Network for Diagnosis of Multiple Sclerosis</i>	140
<i>Stock Wizard: A Machine Learning Model Designed for Day Trading Stock Price Prediction</i>	147

<i>Using Behavioral Forma Mentis Networks to Map Gender Bias in Large Language Models</i>	154
ENGINEERING	158
<i>Creating a Graphene Based Fluorescent Fingerprint Powder</i>	159
<i>Development of Application for Early Detection of Oral Issues Based on Intraoral pH</i>	166
<i>Investigating the Effects of Biodegradable Plastic Breakdown on Arbuscular Mycorrhizal Fungi Colonization</i>	174
<i>O-Patch – Developing an Omega-3 Microneedle Patch for Women with Postpartum Depression</i>	187
ENVIRONMENTAL ENGINEERING	196
<i>Effects of Pharmaceuticals and Ocean Warming on Zooxanthellae Expulsion and Asexual Reproduction in Aiptasia: A Model for Tropical Coral</i>	197
<i>Impact of Heat Events on Urban Heat Island Effects</i>	206
<i>Predicting Summer Heatwaves in the Pacific Northwest Using A Novel Integration of Multiple Machine Learning Models</i>	212
<i>Settlement Preferences and Development of Aurelia Labiata Planula Larvae, on Natural and Anthropogenic Substrates in Current and Climate Change Temperature Scenarios</i>	218
AWARDS/RECOGNITION	230

BEHAVIORAL SOCIAL SCIENCES

Color Me Surprised! Evaluating and Identifying the Impacts of Logo Color Trends on Business Brand Awareness

Nitya Bhartiya

INTRODUCTION

This paper aims to determine whether utilizing logo color trends could be a low-cost method to increase profits for small businesses to large chains, especially after COVID-19's significant impacts on business traffic. Investigating how common and uncommon logo color choices affect consumer perceptions, particularly business brand awareness, across various industries would ultimately provide insights into effective strategies for optimizing logos. Brand awareness is defined as a collection of the brand attitude, attraction, recognition, and recall capabilities of a business (Aslaam, 2006). Prior studies have looked at how colors like red and yellow tend to incite hunger and how such colors are prevalent in the fast-food industry (Labrecque, 2020). The prevalence of these colors in an industry suggests that there may be previously undefined benefits to using those industry-standard colors.

Colors make up to 95% of a consumer's first impression (Faircloth, 2001), which is why studying how colors in business logos lead our brains to make commercial decisions is especially essential to marketing studies. Logos themselves build up a significant amount of a brand's first impression. Companies like Taco Bell spend 20 times more on signs with their logo than on video advertisements (Hendersen and Cote, 1998). The logo is the first thing seen and evaluated even before the product is

even used, and logo color is crucial for capturing consumer attention.

From 2019-2020, during Covid, the eCommerce industry boomed with a 4.4% increase of eCommerce shares in total global retail sales. In April alone of 2020, eCommerce website Etsy, gained 200% of the sellers it had gained in the entire past year. There is an abundant and increasing number of small business owners in need of an effective and adapting online logo that utilizes the advantages of color trends. Advantages that can be seen through the rebranding of ice cream company, Van Leeuwen, increasing profits by 50% relative to the previous year. Good logos, designs, and branding work.

Numerous studies have looked at color psychology amongst logos and in marketing as a whole. One such experiment from the Journal of Business Research (Park et al., 2013) aimed to investigate the impact of brand logos on customer commitment and firm performance. The experiment involved 450 undergraduate students who evaluated logos of 77 corporate brands selected from the Fortune 500 list. Each participant assessed ten brand logos, providing insights into their identification, self-identity/expressiveness, functional benefits, and aesthetic appeal. The study utilized hierarchical linear modeling to analyze the predictors of customer brand commitment. The results revealed that while brand identification through logos did not

significantly influence commitment, the perceived self-identity/expressiveness, functional benefits, and aesthetic appeal of the logos positively correlated with customer commitment. The article's conceptual framework depicted in Figure 1 illustrates how brand logos serve as condensed representations of a brand's marketing efforts, influencing customer perceptions and subsequently impacting firm performance. Brand commitment, defined as customers' willingness to maintain a relationship with the brand, is posited to mediate the effect of strong brand logos on firm performance. In today's

cluttered marketing environment, logos streamline brand identification and facilitate quicker decision-making for customers. Moreover, the study identifies three distinct mechanisms through which brand logos enhance brand commitment and firm performance: facilitating customer self-identity and expressiveness, communicating a brand's functional benefits, and providing aesthetic appeal. These mechanisms highlight the multifaceted role of brand logos in shaping customer-brand relationships and driving business outcomes.

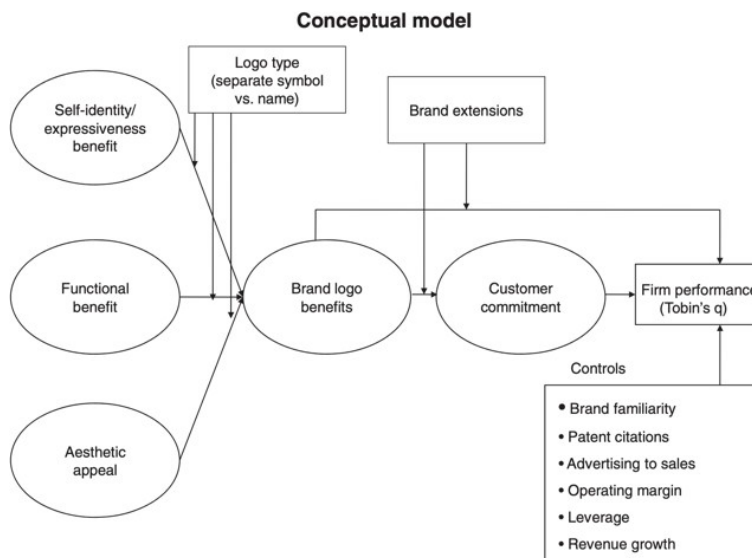


Fig. 1. Conceptual model.

The analysis of brand logo benefits, customer commitment, and firm performance conducted by the authors revealed several significant findings. Utilizing hierarchical linear modeling (HLM) and structural equation modeling (SEM), the study determined the intricate relationships between these variables. Firstly, the HLM analysis demonstrated that while brand identification did not significantly relate to brand commitment ($p > .05$),

the three proposed logo benefits—self-identity/expressiveness, functional, and aesthetic—showed positive associations with commitment, with effect sizes ranging from $\hat{\gamma} = .12$ to $\hat{\gamma} = .18$ ($p < .01$). The significance of these relationships was underscored by the statistical testing, which showed a significant improvement in variance explained when moving from Model 1 to Model 2 ($\Delta \text{Dev} = 1057.19$, $df = 12$, $p < .01$). Additionally, t-tests

comparing logos consisting of brand names versus visual symbols highlighted the superiority of symbols in providing self-identity/expressiveness benefits, functional benefit communication, and aesthetic appeal ($p < .05$). Furthermore, the SEM analysis unveiled the mediating role of customer commitment between brand logo benefits and firm performance, with brand logo benefits significantly impacting both commitment ($\gamma = .78$, $p < .001$) and performance ($\gamma = .31$, $p < .001$). Notably, a high number of brand extensions were found to strengthen the relationships between brand logo benefits and firm performance, with effect sizes ranging from $\gamma = .36$ to $\gamma = .55$ ($p < .001$). Control variables such as revenue growth, advertising-to-sales ratio, operating margin, and leverage also showed significant influences on firm performance ($p < .05$ or $p < .01$). The study's findings underscore the critical role of brand logos in enhancing customer-brand relationships and firm performance, especially when considering the design elements and strategic extensions of logos. However, the study also acknowledges certain limitations, such as the need for longitudinal data analysis to account for changes in brand logos over time and the exploration of other design elements beyond the scope of the current research. Overall, the findings underscored the importance of brand logos, particularly those incorporating visual symbols, in enhancing customer-brand relationships and firm performance in today's competitive marketplace.

Another study done by Labrecque and Milne on the significance of color in marketing, published in the *Journal of the Academy of Marketing Science*, offered insights into the impact of color on consumer

perceptions and brand personality. Through four studies, the authors delve into the relationship between color and brand perceptual measures, drawing on aesthetics, color psychology, and associative learning. Experimentally, they manipulate the hue, saturation, and value of colors across a set of fictitious logos to examine their effects on brand personality dimensions. In Study 1, they focus on hue, keeping saturation and value constant, and find significant relationships between certain colors and brand personality traits. For instance, white, yellow, and pink hues positively influence perceived sincerity, while red, orange, and yellow hues enhance perceived excitement. Moreover, blue and brown hues positively affect perceived competence, while black, purple, and pink hues enhance perceived sophistication. The study involves 279 undergraduate participants, and multivariate regression analyses reveal significant associations between color and brand personality dimensions. Notably, the findings support most hypotheses proposed by the authors, highlighting the importance of color in shaping consumer brand perceptions. The discussion underscores the study's contribution to understanding the interplay between color and brand personality, paving the way for further research in this domain. Experimentally, the study involved carefully controlled manipulations of color variables and utilized multivariate regression analysis to elucidate the relationships between color and brand personality dimensions, underscoring the robustness of the findings. Overall, Labrecque and Milne's research sheds light on the strategic use of color in marketing and emphasizes its role in forming consumer brand perceptions.

An analysis of the results of Labrecque and Milne's study highlights the significant impact of color on brand personality and consumer perceptions. Through paired t-tests on seven dependent variables, including five brand personality dimensions, likability, and familiarity, the study reveals compelling findings regarding the influence of color on brand perceptions. Specifically, the t-tests between color and grayscale versions of the logo show significant changes in the excitement ($t=4.172$, $p=.0001$) and competence ($t=3.546$, $p=.001$) dimensions. Moreover, the combination of logo shape and color significantly alters the ruggedness dimension ($t=-2.249$, $p=.027$). Likability increases significantly with logo shape ($t=2.806$, $p=.006$), color ($t=2.188$, $p=.031$), and their combination ($t=4.634$, $p=.0001$). Moreover, the study explores the impact of specific colors on brand personality dimensions. Separate analyses for predominant colors such as red, blue, and black reveal significant increases in perceived excitement, competence, and sophistication, respectively, between color and grayscale versions of logos. For instance, perceived excitement increases significantly for red logos ($t=2.402$, $p=.024$) and perceived competence increases for blue logos ($t=2.467$, $p=.019$). Additionally, perceived sophistication increases significantly for black logos ($t=2.196$, $p=.035$). These findings suggest that color, when matched with specific brand personality dimensions, enhances consumer perceptions of brand personality.

Another article (Bottomley & Doyle, 2006) delves into the influence of color on brand perception, focusing on the concept of color-product

congruity. Drawing on prior research, the authors emphasize that both colors and products carry connotative meanings, and when their associations align, the chosen color is perceived as more appropriate for the product or brand. This idea of congruity has been prominent in advertising and persuasion literature, where matching the country image with desired product benefits has been shown to enhance purchase intention. Similarly, brand evaluation improves when there is congruity between the brand's emotion benefit proposition and the emotions associated with product usage. An earlier study by Walker et al. (1986) explored the appropriateness of certain typefaces for different professions, finding that appropriateness was determined by the extent to which a typeface shared inherent qualities with the concept it represented. For example, Palatino Italic was judged as feminine and light, fitting for professions like doctors and dentists, while Braggadocio was considered masculine and strong, suitable for builders. Moreover, Lewis and Walker (1989) observed behavioral consequences of typeface perception, showing that congruency between word and typeface connotations led to faster responses in experimental tasks. The article also discusses perceptual fluency, which refers to the ease with which a stimulus is processed. Prior research suggests that increasing perceptual fluency leads to more favorable perceptions of truthfulness, fame, stimulus duration, and likability. Lee and Labroo (2004) distinguish between perceptual and conceptual fluency, where both contribute to a more positive attitude towards a brand. Therefore, the authors infer that congruent color-product combinations will be processed more

fluently, leading to higher likability and perceived appropriateness compared to incongruent combinations.

Through two experiments, it investigates how color choices impact the image of brands. Experiment 1 assesses color appropriateness for functional and sensory-social products, while Experiment 2 examines the effects of red and blue colors on brands positioned in both categories. The study hypothesizes that the suitability of a color for a brand is determined by its congruence with the product type. Thirty-five participants, including teachers and administrative staff, evaluated 96 logos representing various products and colors. Each logo displayed the brand name in a distinct color, with products classified as functional or sensory-social. The methodology involved individual testing sessions and booklet presentations of logos for evaluation, followed by manipulation checks to ensure participants' understanding. The findings support the hypothesis that color appropriateness is indeed influenced by the congruity between the product type and the color choice, indicating an importance of the industry behind the product and the logo color.

The cited literature suggests that logo color trends hold significant implications for consumer perceptions and brand performance. Studies by Park et al. (2013) and Labrecque and Milne underline the positive correlation between logo attributes such as self-identity, functional benefits, and aesthetic appeal with consumer commitment, thereby impacting business performance and traffic. Additionally, Bottomley and Doyle's research

emphasizes the importance of color-product congruity in enhancing brand perception.

With these prior studies serving as significant background research and additional preceding investigations on the association between colors and their effects on consumers, this novel study aimed to build upon existing knowledge by investigating a correlational relationship between usage of industry-based color trends in logos and overall business brand awareness, defined as brand attitude, attraction, recognition, and recall. It is hypothesized that logo color choices with industry standards and consumer expectations through logo color trends will lead to higher businesses evaluations and consequently, traffic, particularly important in the post-COVID-19 era of commerce. The hypothesized results of this study were that using identified industry-standard (common) colors will lead to increased brand attitude, attraction, recognition, and recall ($r \geq 0.2$) relative to using identified uncommon colors.

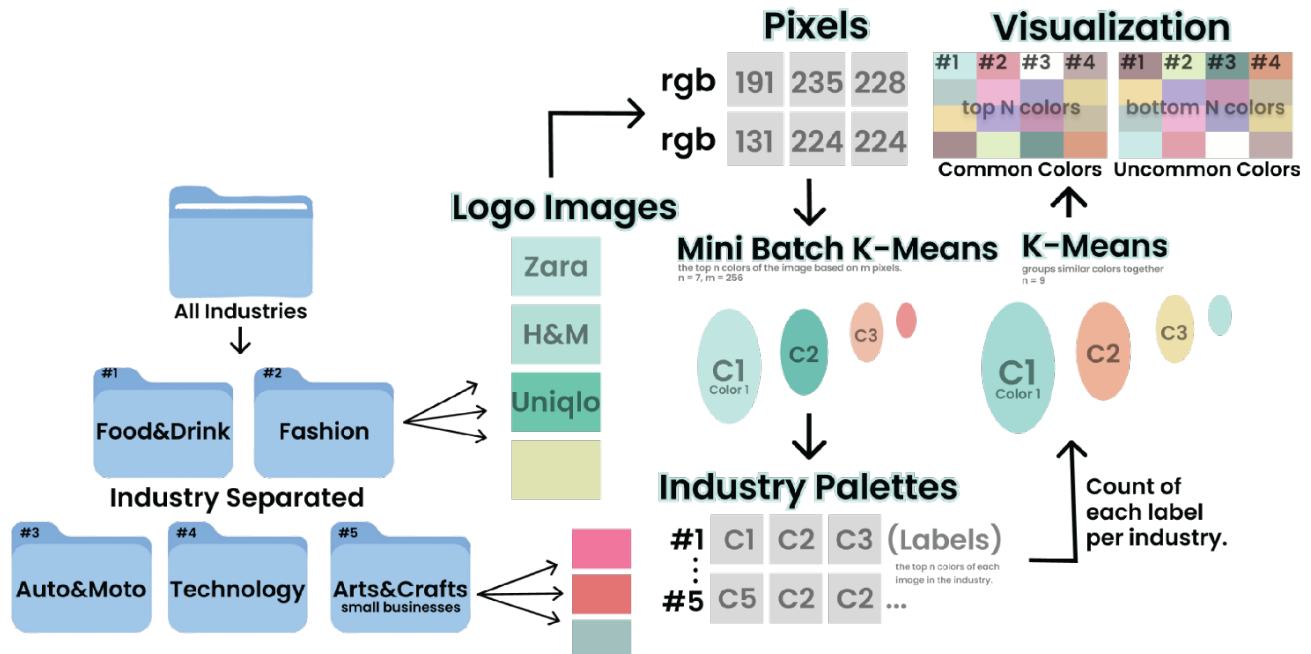
METHODOLOGY

The methodology consisted of data collection and preprocessing for a Python script to determine the industry-standard colors. Then, those colors were inputted to create 20 logos using text-to-image AI and evaluated across different brand metrics through a survey. Data collection was conducted using an online website called 1000logos.com, with over 500 logos from five sectors—food and drink, fashion, auto and moto, technology, and arts and crafts—being collected. Adobe Photoshop was utilized for data preprocessing, involving resizing the logos to 1080 by 1080 pixels and exporting them as PNG files with

RGB color mode. This step ensured that the dataset was sufficiently large and diverse to provide insights into color usage within each industry. Following data collection, a Python script leveraging OpenCV libraries was developed to determine

common and uncommon colors present in the logos through mini batch k-means clustering. Figure 1 below depicts a flowchart detailing the Python script.

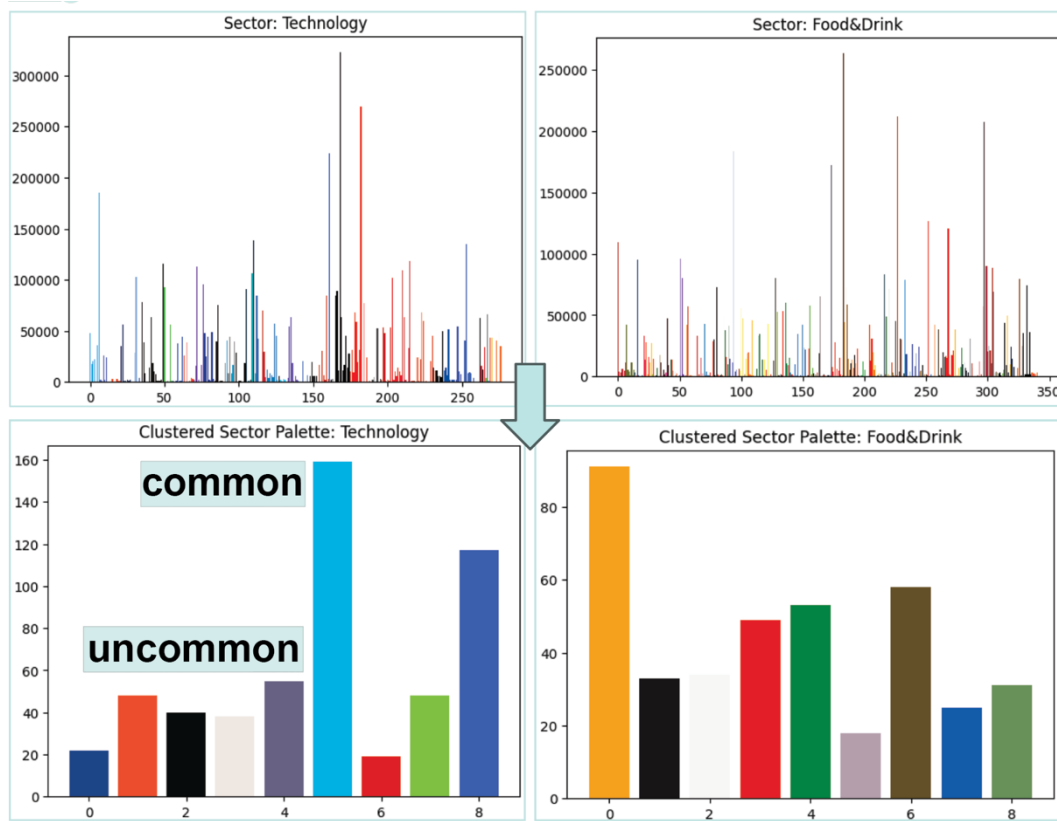
Figure 1. Python Script Algorithm Flowchart



A Python script leveraging mini batch k-means clustering was utilized to identify common and uncommon colors present in the logos. For each industry, each image was converted into pixels containing RGB color data. The most prominent 7 colors from each image were found with Mini Batch K-Means and formed the industry palettes. To

narrow down the colors extracted in the, there was a second round of clustering to retrieve 9 colors total per industry, the highest being the most common color. This approach enabled the extraction of key color palettes representative of each industry, laying the groundwork for subsequent logo creation and is summarized in Figure 2 below.

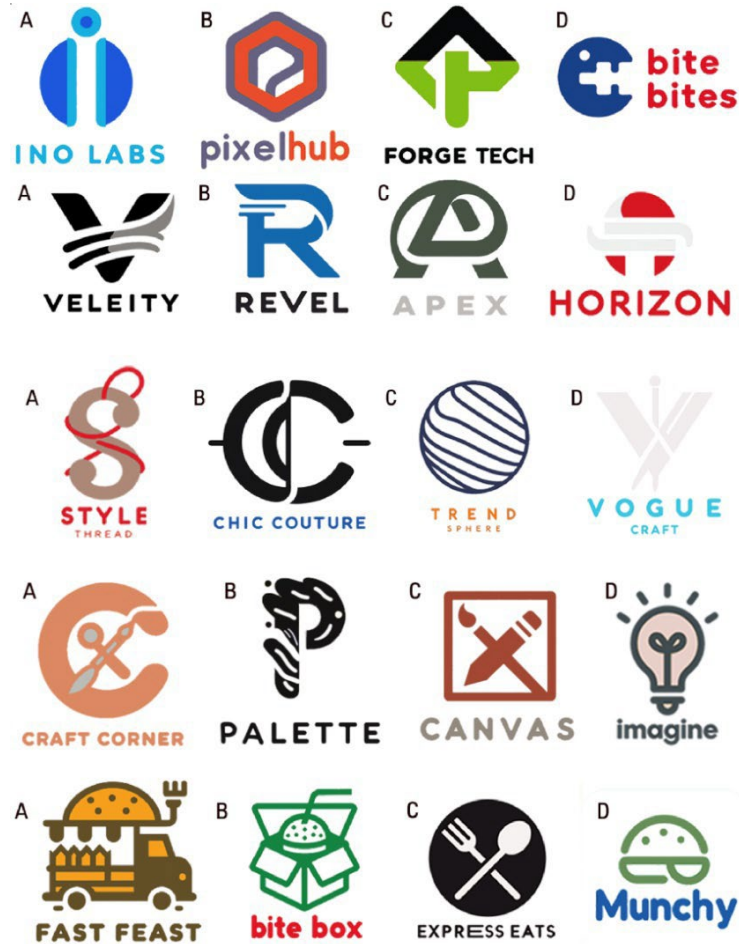
Figure 2. Python Chart Visualizations



The procedure also involved the creation of logos using AI technologies. Four logos were generated using the colors depicted in Figure 2 for each industry, with two using the top four common colors and two using the top four uncommon colors identified earlier. The logos were standardized in size and color, featuring both an icon and brand name. To create 20 logos, Microsoft Copilot including GPT-4 and DALL-E 3 were used. The usage of AI was necessary to prevent personal experimenter bias and to better standardize the creation of all logos. The first step involved asking Microsoft Copilot to generate logo names. This gave

a name to input for the second prompt in which DALL-E, a text-to-image AI, was asked to create a simple minimal two-color logo given the business name and industry. A prior approach was taken through providing DALL-E with a specific RGB color code. However, DALL-E used the semantics behind a logo's name instead to inform its color choices, causing a need to recolor the logos appropriately in Adobe Photoshop. Figure 3 below shows all 20 logos corroborated across their respective industries. Subsequently, a survey sampling procedure was implemented to gather participant evaluations of these created 20 logos.

Figure 3. All 20 logos generated with DALL-E



Sixty-four adolescents (aged 13-70), composed of a near equal female to male ratio, were obtained from 3 school districts in Washington state—Lake Washington School District, Bellevue School District, and Seattle Public Schools—via submitting a public records request at those district websites and distributing flyers. 33 participants were selected using a random number generator to participate in the experiment and assigned numbers in random order. The survey, administered through

Tally, consisted of 5 questions per industry, totaling 25 questions, all emphasizing the logo colors to assess logos in attitude, attraction, recall, and recognition. Demographic information including age, gender, and race was collected to provide context for the responses. Data analysis involved correlating participant ratings with logo characteristics and demographic factors to identify trends and associations.

Data Table 1. *Participant Logo Brand Metric Evaluations*

Participant #	Logo Industry	Common or Uncommon Colors Used (0-1)	Brand Attention (1-4)	Brand Attitude (1-4)	Brand Recall (0-1)	Brand Recognition (0-1)
1	Fashion	0	2	2	1	0

RESULTS

The hypothesized results of this study were that using identified industry-standard (common) colors will lead to increased brand attitude, attraction, recognition, and recall ($r \geq 0.2$) relative to using identified uncommon colors. The findings of the study confirmed this hypothesis, revealing statistically significant correlations between the usage of common colors and enhanced brand evaluation metrics ($r \geq 0.4$). Specifically, the analysis demonstrated that logos employing common colors elicited higher mean scores for brand attention, value, attitude, and recall compared to logos utilizing uncommon colors. Particularly strong results were shown in the brand recall metric with a Pearson's Correlation Coefficient of 0.61 and a recall count of 9 times more on average when using common colors relation to using uncommon colors ($p < 0.0001$). These results provide support for the notion that aligning logo colors with industry trends can effectively enhance brand awareness and favorability, particularly in a post-pandemic business

landscape with millions of more small business owners where consumer engagement is paramount.

Twenty-five paired T-tests were conducted for brand attitude, attraction, recognition, and recall showed statistically significant differences between logos using common and uncommon colors, as evidenced by p-values below 0.05 for all metrics measured (bolded in Figure 4). Logos utilizing common colors were associated with higher mean scores for brand attention, value, attitude, and recall compared to logos using uncommon colors. Specifically, the auto & moto industry demonstrated the most significant impact, ($p < 0.01$), while the fashion industry showed less significant effects. Results indicated statistically significant differences between logos utilizing common and uncommon colors. Specifically, logos using common colors garnered higher mean scores for brand attention, value, attitude, and recall. Pearson's correlation coefficient analysis revealed moderate to strong positive correlations of 0.46, 0.47, 0.4, and 0.61, respectively, between the usage of common or uncommon colors and brand evaluation metrics.

Figure 4. Brand Metrics Consolidated

ALL	Attitude		Attraction		Recognition		Recall	
	Common	Uncommon	Common	Uncommon	Common	Uncommon	Common	Uncommon
Mean	3.4879	3.2788	3.4303	3.2636	8.7879	7.3636	6.6667	3.7576
t Stat	2.0942		1.9451		4.0583		7.4567	
P(T<=t)	0.0221		0.0303		0.0001486		8.68E-09	
MOST	Auto&Moto		Auto&Moto		Auto&Moto		Fast Food	
	Common	Uncommon	Common	Uncommon	Common	Uncommon	Common	Uncommon
Mean	3.5455	3.2273	3.4697	3.1212	1.8182	1.3030	1.2727	0.5455
t Stat	2.2090		2.1697		3.7208		4.9774	
P(T<=t)	0.0172		0.0188		0.00038		1.061E-05	
LEAST	Fashion		Fashion		Technology		Arts&Crafts	
	Common	Uncommon	Common	Uncommon	Common	Uncommon	Common	Uncommon
Mean	3.2121	3.1515	3.2121	3.1818	1.8182	1.6667	1.3333	0.8182
t Stat	0.3409		-0.3782		1.3047		3.0441	
P(T<=t)	0.3677		0.3539		0.1007		0.00232	

CONCLUSIONS

The study supported the hypothesis by revealing statistically significant associations between logo color usage and brand evaluation metrics. Pearson's correlation coefficient analysis indicated that businesses could enhance brand awareness, favorability, and familiarity by aligning logo colors with industry trends. The findings suggest that utilizing common colors in logos could serve as a low-cost method for small businesses to increase brand traffic and profitability.

Limitations were identified, including the relatively small sample size of 33 participants from three states, which may have introduced variations in the data. Additionally, the time-consuming nature of organizing survey data and the potential for bias in DALL-E generated logos were recognized. Participants may also have been influenced by factors other than color attributes when evaluating logos. To address these limitations, future research could focus on gathering more participant responses across a wider demographic to strengthen the results.

Exploring logo trends across additional industries and factors, such as typography differences, could further validate the findings and mitigate bias. Additionally, employing ANOVA tests to consider factors such as ethnicity, gender, and color wavelengths could provide deeper insights. Developing an app to streamline the process of logo analysis for businesses could also be a practical application of the research findings while also employing a survey that compares two logos of the same icon while changing just the colors. Continued research in this area has the potential to empower businesses of all sizes to leverage color psychology effectively in their branding strategies.

BIBLIOGRAPHY

- Aslam, M. M. (2006). Are You Selling the Right Colour? A Cross-cultural Review of Colour as a Marketing Cue. *Journal of Marketing Communications*, 12(1), 15–30. <https://doi.org/10.1080/13527260500247827>
- Bottomley, P. A., & Doyle, J. R. (2006). The interactive effects of colors and products on perceptions of brand logo appropriateness. *Marketing Theory*, 6(1), 63–83. <https://doi.org/10.1177/1470593106061263>
- Labrecque, L. I. (2020). Color research in marketing: Theoretical and technical considerations for conducting rigorous and impactful color research. *Psychology & Marketing*, 37(7), 855–863. <https://doi.org/10.1002/mar.21359>
- Labrecque, L. I., & Milne, G. R. (2012). Exciting red and competent blue: the importance of color in marketing. *Journal of the Academy of Marketing Science*, 40(5), 711–727. <https://doi.org/10.1007/s11747-010-0245-y>
- Lee, A. Y., & Labroo, A. A. (2004). The Effect of Conceptual and Perceptual Fluency on Brand Evaluation. *Journal of Marketing Research*, 41(2), 151–165. <https://doi.org/10.1509/jmkr.41.2.151.28665>
- Lewis, C., & Walker, P. (1989). Typographic influences on reading. *British Journal of Psychology*, 80(2), 241–257. <https://doi.org/10.1111/j.2044-8295.1989.tb02317.x>
- Park, C. W., Eisingerich, A. B., Pol, G., & Park, J. W. (2013). The role of brand logos in firm performance. *Journal of Business Research*, 66(2), 180–187. <https://doi.org/10.1016/j.jbusres.2012.07.011>
- Walker, P., Smith, S., & Livingston, A. (1986). Predicting the appropriateness of a typeface on the basis of its multi-modal features. *Information Design Journal*, 5(1), 29–42. <https://doi.org/10.1075/idj.5.1.02wal>

Competitive Learning & Cognitive Resilience: Analyzing Peers' Influence on Planning and Problem-Solving Skills

Priyanka Kudallur

INTRODUCTION

This project seeks to investigate the impact of peer pressure on cognitive performance of high school students, specifically through an assessment of executive planning and problem-solving skills using a digital version of the Tower of London test. The Tower of London test is an existing neuropsychological assessment of individuals' planning, problem-solving, and cognitive flexibility. Participants rearrange colored beads on pegs to match a target pattern within five minutes, recording the number of sets completed and moves taken per set. By integrating this metric with electroencephalography (EEG), this investigation aims to utilize behavioral and neuroscientific data to elucidate the adverse effects of a competitive environment on high school students' performance.

Anxiety and depression are among the leading causes of illness and disability in adolescents, and the primary effector of these mental health issues is academics: 61% of American teens report pressure to get good grades (Horowitz, J. M.). While these pressures may be intrinsic or familial, the most pervasive is peer pressure, which can motivate students while simultaneously subjecting them to severe stress. This phenomenon is called competition-based learning. Prior studies, such as from the Journal of Youth and Adolescence, have established the exacerbation of risky behaviors in adolescents due

to peer pressure, wherein pursuit of conformity drove increased skipped classes, alcohol consumption, and attitudes toward romantic partners (Santor, D. A). However, the effect of peer pressure on academic performance, which involves higher-order cognitive processes, is still contested with varying scientific results. Uncovering the intricate relationship between competition-based learning (CBL) and its hindrance or improvement of short-term planning and problem-solving skills will yield key insight into whether a competitive learning environment is conducive to the education system or a contributor to adolescent stress and anxiety.

Prior studies have explored competition-based learning environments and whether they improve student performance in the long term, however the results tend to vary on a student-by-student basis and their preference for competition and gamification of learning. This relates to Social Comparison Theory (SCT), which describes how comparison of one's performance to others shapes our self-image and evaluation of abilities. According to the American Psychology Association, comparison can either lead to the contrast effect wherein there is a resulting negative change in self-esteem or the assimilation effect the proactive response which instead promotes inspiration in the individual (American Psychology Association). In order to determine whether

education practices which reinforce student comparison are conducive to learning, the prevalence of the negative versus positive response amongst adolescents must be determined.

One study sought to incorporate EEG data into understanding this phenomenon, using frontal alpha and frontal beta power as a measure of stress. However, the use of an ineffective social-evaluative threat (a video camera was used to simulate social pressures) contributed to the study ultimately finding no change in EEG power (Ehrhardt, N. M.). Unlike prior studies which have been conducted longitudinally and/or in a digital environment, this study will use a novel method which is better reflective of test taking-peer pressure. Through the administration of the Tower of London test in two environments, this study stimulates self-comparison and competitiveness with the physical component of other students competing to complete the task first. Existing literature affirms the relation between social decision-making and worsened mental health amongst adolescents – at this age, depression detection rates range from 20 to 44% (Hu, Y.). Based on this information, the effects of self-comparison can be implied as largely negative and stress-inducing for high school students. By measuring brain activity of students in an isolated setting carrying out the puzzle task versus in a competitive environment, this study could reveal heightened mental stress in individuals when competing to be the first ones finished. Harmful self-comparison to peers would be characterized by greater sensitivity of the frontal lobe region under the electroencephalography equipment.

METHODOLOGY

Hypothesis Testing

The expected result of worsened planning and problem-solving performance in the group environment will be measured based on a reduction in completed *Tower of London* test sets completed by students in the presence of peers. These results will be analyzed alongside the EEG-collected data of brain wave propagation in the independent versus group environments, to pinpoint the neural basis for social comparison and its potentially detrimental effect on planning and problem-solving for adolescents.

Participants and Equipment

The study will involve a minimum of 25 human participants. Electroencephalography (EEG) data will be collected using the OpenBCI Ultracortex Biosensing Headset, a device purchased by Tesla STEM High School for academic use. EEG signal analysis will be conducted using the EEGLAB toolbox, an open-source extension within the MATLAB programming environment, downloadable from <https://sccn.ucsd.edu/eeglab/index.php>. Additional equipment includes a battery-powered Microsoft Windows Surface laptop for data collection, Ten20 Conductive Paste (a medical-grade adhesive) for securing EEG electrodes, a stopwatch, and five school-administered laptops (DELL Latitude 3190 2-in-1) with Internet access for the Tower of London application. The study will also utilize Microsoft Forms for questionnaires, the Calendly scheduling website, Microsoft Excel for data parsing, an online random number generator, and standard furniture.

Tower of London Test

As mentioned above, the Tower of London (TOL) test is an existing neuropsychological assessment, currently administered physically for the most part. In order to collect the necessary data for this study, the TOL test required a built-in five-minute timer as well as the capacity to measure the total games completed and moves taken (clicks of a bead from one spot to another) within the provided time frame. Given that there lacked such an open-source website or software online, a new version of the TOL test was created in Java for this study. The test was coded in the Java integrated development environment called jGrasp and run on students' laptops through a transferable program file. Two versions of the TOL program were created – one with and one without leaderboard functionality. As shown in *Figure 4*, the group environment features a leaderboard within the sidebar of the open testing window. As the game progresses throughout the five minutes, the leaderboard live updates as each student completes another set. This added feature effectively simulates the pressure of self-comparison and competition during the group testing round.

Sampling Procedure

To identify high school participants in the Lake Washington School District (LWSD), the study will use the official school district directory

obtained through Microsoft Outlook mailing lists. After compiling a filtered list of high schools within LWSD, assigning numbers, and calculating the total number of students, the study's optimal sample size (n) is calculated by dividing the total number of students by 5. Random sampling will be employed to select 20% of high school students, starting at a random point, selecting every fifth student's email address, and deleting other values. Selected students will receive email and participation surveys, including information about the study's purpose and data use. Additionally, flyers with a study description and a QR code for confirmation will be posted in high schools. Follow-up communication will involve obtaining informed consent digitally, with completion serving as confirmation of participation.



Figure 1: *Flyers to Distribute Around Local High Schools*

Research Study: Preliminary Questionnaire

Rate the following questions on a scale of 1 to 5 based on how well they apply to you. An example scale is modeled here:

1
2
3
4
5

strongly disagree *strongly agree*

1. I often compare my academic abilities to those of my peers.	1	2	3	4	5
2. I frequently evaluate my accomplishments in comparison to others.	1	2	3	4	5
3. I often feel the need to outperform my peers in academic tasks.	1	2	3	4	5
4. I am aware of how I rank academically among my peers.	1	2	3	4	5
5. I often find myself comparing my academic progress to others.	1	2	3	4	5
6. I strive to do better academically than others in my class.	1	2	3	4	5
7. I enjoy challenging myself academically to reach higher levels of achievement.	1	2	3	4	5
8. I feel motivated to excel academically to improve myself.	1	2	3	4	5
9. I am open to learning new things even if it means competing with others.	1	2	3	4	5
10. I see competition as an opportunity to grow and improve my skills.	1	2	3	4	5
11. I am motivated to do well in my academic studies.	1	2	3	4	5
12. I find academic challenges enjoyable and stimulating.	1	2	3	4	5
13. I am committed to putting effort into achieving academic success.	1	2	3	4	5
14. I am driven to achieve high grades in my classes.	1	2	3	4	5
15. I am enthusiastic about learning new things in my academic subjects.	1	2	3	4	5

Figure 2: Document Version of the 15-Item Preliminary Questionnaire

Statements:

- I often compare my academic abilities to those of my peers.
- I frequently evaluate my accomplishments in comparison to others.
- I often feel the need to outperform my peers in academic tasks.
- I am aware of how I rank academically among my peers.
- I often find myself comparing my academic progress to others.
- I strive to do better academically than others in my class.
- I enjoy challenging myself academically to reach higher levels of achievement.
- I feel motivated to excel academically to improve myself.
- I am open to learning new things even if it means competing with others.
- I see competition as an opportunity to grow and improve my skills.
- I am motivated to do well in my academic studies.
- I find academic challenges enjoyable and stimulating.
- I am committed to putting effort into achieving academic success.
- I am driven to achieve high grades in my classes.
- I am enthusiastic about learning new things in my academic subjects.

3 Subscale Factors:

- Academic Motivation
- Self-Comparison
- Competitiveness and Development

Figure 3: Questionnaire Statements Divided by Subscale Factors

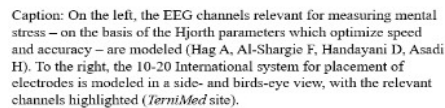
Caption: The fifteen statements (evaluated on a scale of 1 to 5) are categorized into three distinct factors, for which the responses will be grouped together when calculated cumulative subscales for each participant.

Experimental Procedure – Individual Testing

For round 1 data collection, participants will be requested to provide their individual availability through a Calendly link with an interval of approximately ten days. In a public library space, participants will be individually greeted, seated at a

table with their school laptops, and guided to log into the Tower of London application. EEG equipment setup will involve the application of Ten20 Conductive Paste and adherence to the international 10-20 system for electrode placement, ensuring optimal signal quality and electrical safety.

completed puzzle sets, moves taken, and the EEG data stream link for each participant, following a specific naming convention.



Model 1: Model of EEG Equipment and Placement

In more detail, the electrode placement will adhere to the international 10-20 system, with relevant electrodes (AF3, FC5, F8, FP1, AF4, P7, Fp2, F7) attached. Given the study's compatibility with a 32-electrode EEG, specialized channels (AF3, AF4, FC5) will be positioned using nearby channels from the 10-20 system as a reference. The secure attachment of electrodes to the EEG recording device will ensure a stable and strong connection without causing discomfort. Impedance levels will be continuously monitored to guarantee optimal signal quality, and any adverse skin reactions will prompt an immediate halt to the experiment for the affected participant.

Experimental Procedure – Group Testing

Randomly assign participants to groups of five for round 2, determining the number of mini

groups based on the quotient of n divided by 5.

Instruct students in groups, set up tables and chairs, and explain the competitive round with a live leaderboard. Randomly choose one student for EEG data collection in each trial. Countdown to start the puzzles, initiate the stopwatch, and collect EEG data simultaneously for five minutes. Record puzzle sets completed, moves taken, final leaderboard ranking, and EEG data stream link for the student wearing the EEG, using a specific naming convention. Enter "n/a" for other group members. Inform students that only one round of competitive testing is needed and that the five-minute puzzle period will not be repeated for each student wearing the EEG due to logistical constraints.

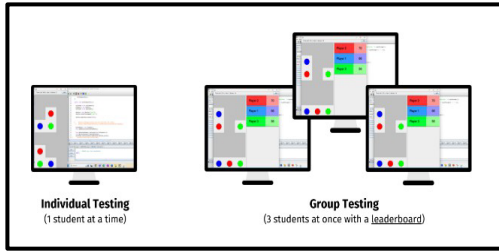
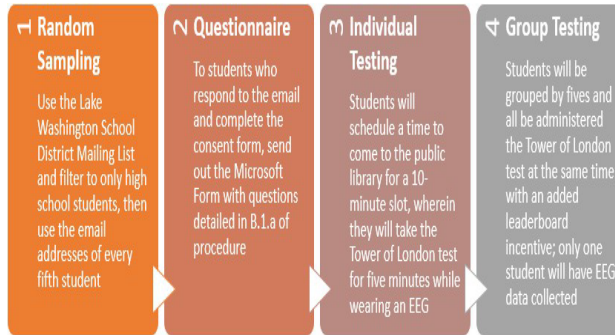


Figure 4: *Model of Individual vs. Group Tower of London Testing*

Caption: The Tower of London program is run in the jGrasp integrated development environment. Group testing is different from individual testing through the addition of the leaderboard in the program, which updates students' scores live as they progress.

Flowchart of Procedure



Caption: As a human-based study, this experiment involves a two-part procedure split into sampling and experimental. (1) The first step of the procedure is random sampling to obtain at least 25 participants, who must complete the consent form after which they will be (2) sent the initial questionnaire and a meeting scheduling link. From there, the experiment moves into the (3) individual testing and (3) group testing of the *Tower of London* application as detailed in the procedure above.

CONCLUSION

Data

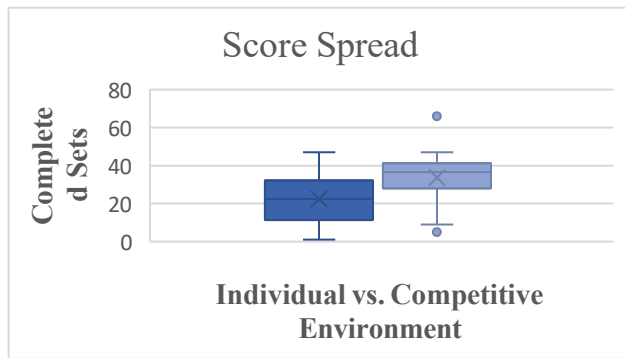
Table 1: *Filtered Data Table of Completed Scores Comparison*

Participant #	# Completed Sets (Individual)	# Completed Sets (Group)
21	12	22
25	9	32
19	1	5
27	14	41
8	26	38
12	39	47
23	31	37
28	21	30
11	23	30
34	41	24
4	3	9
32	36	36
3	48	35
22	31	42
31	22	38
10	52	36
14	47	66
33	56	23
AVERAGE_RAW	16.75154856	13.71667084
AVERAGE	22.5	33.78571429

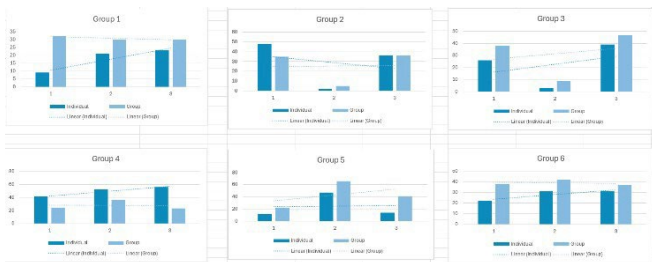
Table 2: *Questionnaire Subscale Scores*

ID	Self-Comparison	Competitiveness & Development	Academic Motivation
3	19	25	25
4	24	17	21
8	15	17	16
10	20	23	23
11	23	21	25
12	19	16	17
14	18	9	8
19	15	20	22
21	20	20	20
22	22	20	23
23	11	14	19
25	20	23	25
27	22	20	20
28	21	24	24
31	20	25	25
32	18	16	19
33	23	19	22
34	23	21	24
AVERAGE	19.61111111	19.44444444	21

Graph 1: Score Comparisons for Individual Environment-First Participants



Graphs 2: Score Comparisons by Competition Group



RESULTS – Tower of London Analysis

Filtered comparison of the total number of sets completed in the individual and competitive environment yielded an average 66.596% increase, with the paired t-test establishing statistical significance ($p=0.0000861$). After removing the four outlier values (highlighted yellow in *Table 2*), the average completed sets in the individual environment was 22.5 and about 33.8 in the group environment. Such results would support the acceptance of the alternate hypothesis, that competition-based learning improves students' problem-solving skills.

RESULTS – Sample Analysis

The improvement in performance of individuals in this study refutes the original

hypothesis that the competitive environment would worsen performance. Deeper analysis of the preliminary questionnaire which participants completed provides insight into why students may have performed better while undergoing self-comparison to their peers. *Table 2* reveals the calculated sub-scores for each of the three categories – self-comparison, competitiveness and development, and academic motivation – for each participant. These sub-scores were computed by adding up the individual scores for each of the five relevant statements (refer to key in *Figure 3*). The average cumulative scores for these three categories were approximately 19.6, 19.4, and 21.0, respectively. Given that the highest possible cumulative score for any given category was 25 (5 questions ranked on a 5-point scale), these averages indicate that the sample of 18 participants in this study was self-reported to be highly competitive. Such results reinforce the idea that students' response to competitive pressure has a strong degree of student-to-student variation, which should be considered when making larger generalizations about optimal school environments.

IMPLICATIONS

This study reveals how some students respond to competitive stress, this sample particularly showing an improved response in short-term problem-solving. Looked at in isolation, these results would suggest that educational environments should encourage a healthy amount of student-to-student comparison to stimulate increased motivation to excel academically. However, the generalizability of this study is limited by the

sample, wherein students self-reported to be highly competitive and prone to self-comparison.

LIMITATIONS

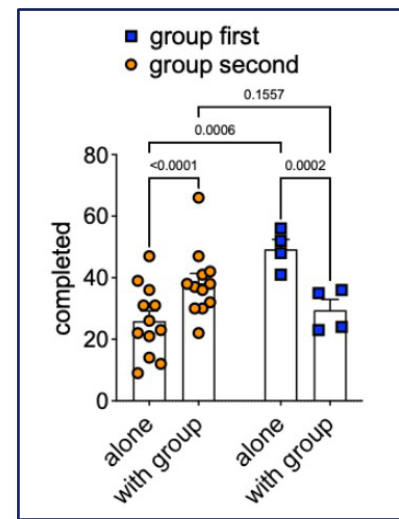
As mentioned above, a key limitation of this study is the sample size and representation. Future research should encompass a greater participant pool by expanding the sampling procedure. Another experimental limitation uncovered during the analysis portion of the study was the administration of the individual versus competitive environments. Fourteen of the eighteen participants first took the Tower of London test individually and then the group environment (as planned), while four students completed the group environment prior to the individual environment (based on scheduling needs). These four students are highlighted as outliers in *Table 1*. Only upon exclusion of these outliers, was a clear effect of experimental environment seen when conducting the t-test between the individual and group testing scores. This raises the question of whether experiential learning – wherein students grow more comfortable with the testing apparatus – could have played a role in the group environment improvement.

FURTHER ANALYSIS

Further analysis was carried out to consider an added variable of testing administration order (individual versus group first) and its effect on performance. A Two-Way Mixed Measures ANOVA was conducted to consider the within-subjects factor (individual vs. group environment) along with the between-subjects factor (individual test first vs. group test first). Given the statistical significance of the interaction effect ($p < 0.05$) between both factors, post-analysis was carried out using Tukey's

test. Tukey's revealed a key, novel finding that there was a lack of statistical significance between competitive test scores regardless of whether prior testing was done. This alludes to the psychology concept of social facilitation, wherein the group average sets the precedent for individuals' performance. Further, this phenomenon would suggest that the potential experiential learning with TOL does not have a linear effect of improvement on students' performance.

Graph 3: Data Grouped by Test Environment and Test Order, Graphed with Standard Error Bars



Tukey's Multiple Comparison Test	Mean Diff.	q	Significant? P < 0.05?
A-Individual vs A-Group	-12.33	4.113	Yes
A-Individual vs B-Individual	-23.33	5.503	Yes
A-Individual vs B-Group	-3.583	0.8450	No
A-Group vs B-Individual	-11.00	2.594	No
A-Group vs B-Group	8.750	2.063	No
B-Individual vs B-Group	19.75	3.803	No

Table 4: Tukey's Test Results

Results from Tukey's test uncover that regardless of whether participants had prior experience, their scores in the group setting were statistically indistinguishable.

FUTURE RESEARCH

Further research should be conducted to include a more diverse high school student

population. Within this sample, participants self-reported high competitiveness and motivation in preliminary questionnaire – indicating that the results from this study may not be generalizable to a larger high school population. Sampling can be improved through in person advertising and outreach to schools in a wider region. With regards to experimental procedure, studies which draw from this experiment should employ a split-design wherein half of the participants experience the competitive environment first and the other half, individual first. The conclusion of this study also suggested an improved response to stress in the measured sample. However, when considering

broader education practices, it is critical to understand whether academic improvement may come at the cost of students' mental health. Analysis of collected electroencephalography data would yield key findings into whether problem-solving regions of the brain (i.e., frontal lobe) or emotional centers (i.e., amygdala) were more strongly exercised in the individual versus competitive environment. Although this EEG data was collected throughout the experiment, a limiting factor of this study was the learning curve with data analysis of the EEG data streams in the provided time. Next steps include integrating the change in beta power from EEG data to support drawn conclusions.

BIBLIOGRAPHY

American Psychology Association. (n.d.). *Social Comparison Theory*. Retrieved from <https://dictionary.apa.org/social-comparison-theory>

Balaji, N., Murthy, P., Kumar, D., & Chaudhury, S. (2019). Perceived stress, anxiety, and coping states in medical and engineering students during examinations. *Industry Psychiatry Journal*, 28(1). https://doi.org/10.4103/ipj.ipj_70_18

Ehrhardt, N. M., Fietz, J., Kopf-Beck, J., Kappelmann, N., & Brem, A.-K. (2022). Separating EEG correlates of stress: Cognitive effort, time pressure, and social-evaluative threat. *European Journal of Neuroscience*, 55(9-10), 2464–2473. <https://doi.org/10.1111/ejn.15211>

Hag A, Al-Shargie F, Handayani D, Asadi H. (2023, September 18). Mental Stress Classification Based on Selected Electroencephalography Channels Using Correlation Coefficient of Hjorth Parameters. *Brain Sciences*, 13(9):1340. <https://doi.org/10.3390/brainsci13091340>

Horowitz, J. M. (2019, February 20). *Most U.S. teens see anxiety and depression as a major problem among their peers*. Pew Research Center. <https://www.pewresearch.org/social-trends/2019/02/20/most-u-s-teens-see-anxiety-and-depression-as-a-major-problem-among-their-peers/>

Hu, Y., Zhou, M., Shao, Y., Wei, J., Li, Z., Xu, S., Maguire, P., & Wang, D. (2021) The effects of social comparison and depressive mood on adolescent social decision-making. *BMC Psychiatry* 21, 3. <https://doi.org/10.1186/s12888-020-02928-y> Santor, D. A, Messervey, D., & Kusumakar, V. (2000). Measuring peer pressure, and conformity in Adolescent boys and girls: Predicting school performance, sexual attitudes, and substance abuse. *Journal of Youth and Adolescence*, 29(2), 163-182. <https://doi.org/10.1023/A:1005152515264>

Šimek, D., Grum, D. K. (2011). Competitiveness and Motivation for Education in Self-Handicapping. *Studia Psychologica*, 65(3), 83-96. <https://doi.org/10.15804/tner.2015.39.1.16>

Yang, Q., Chang, S., Hwang, G., & Zou, D. (2020, January 7). Balancing cognitive complexity and gaming level: Effects of a cognitive complexity-based competition game on EFL students' English vocabulary learning performance, anxiety and behaviors. *Computers and Education*, 148(3):103808. <https://doi.org/10.1016/j.compedu.2020.103808>

Eliminating Cross-Race Bias in Witness Scenarios through Childhood Exposure

Alex Zhang

INTRODUCTION

This paper aims to report on patterns and correlation between childhood exposure to various origins and the aptitude of individuals in face identification, akin to that of a witness, shedding light on the interplay between early racial experiences and the ability to have accurate face recognition. The story of Ronald Cotton is a poignant illustration of the profound impact of cross-race bias on wrongful convictions. In the 1980s, Jennifer Thompson, a young college student, was the victim of a brutal sexual assault. In the aftermath, she identified Ronald Cotton, an African American man, as her assailant. Despite Cotton's protests of innocence, he was convicted largely based on Thompson's eyewitness testimony. Years later, DNA testing became a pivotal tool in reexamining criminal cases. In 1995, new testing revealed that Ronald Cotton was not the perpetrator, referenced from *State v. Cotton*, 99 N.C. App. 615(1990). Instead, another man, Bobby Poole, was identified as the actual assailant. The revelation not only exonerated Cotton but also exposed the fallibility of eyewitness identification, particularly when it comes to cross-race situations. Thompson's misidentification of Cotton highlighted the phenomenon known as *cross-race bias* or *own-race bias*. This psychological phenomenon suggests that individuals tend to be more accurate in identifying faces of their own racial or ethnic group and are

more prone to errors when identifying faces from other racial backgrounds. The story of Ronald Cotton underscores the need for a critical examination of the role cross-race bias plays in the criminal justice system. It prompts a deeper inquiry into the factors influencing eyewitness testimony and the potential consequences of relying solely on identification in legal proceedings. The case serves as a powerful narrative in the ongoing dialogue about the need for reform and increased awareness of the complexities surrounding eyewitness identification, particularly in diverse and multicultural societies. From the Innocence Project (2024) alone, 249 people have been exonerated so far, with an astounding 63% having to do with reasons relating to eyewitness misidentification. Furthermore, general studies relating to cross-race bias have revealed that individuals are 1.56 times more likely to falsely identify a face of another race. Especially in court for sexual assault cases, eyewitness error is much higher for Black defendants, with 74% of cases compared to the 53% of cases with a white defendant (Gross, 2022). This highlights the need to identify an objective basis behind why this occurs in the criminal justice system.

In developmental psychology, it is already established that preferences for races are related to exposure to other races during early years of

development. By college, which conventionally begins when one is 18 years old, cross-race bias has already been established to be present in white, college level individuals (Davis et al., 2016; Wright et al., 2003). While there have been studies showing the presence of cross-race bias in court, like the effect that childhood development affects adult abilities and preferences, one might hypothesize that childhood development in cross-race exposure may lead to inaccuracies in eyewitness identification due to the cross-race effect. In a study conducted on cross-race bias in primarily elementary school, analysis of friendships helped lead the researchers to conduct a study on racial preferences in these young children, showing a preference for own-race hypothetical play partners based on their friendships in first and third grade. This demonstrates an aspect of childhood exposure and its impact on the early racial preferences in life (Gaias et al., 2018). Moving onto the situational aspects that have been studied previously that investigate the nature side of development psychology, exposure to the cross-race bias as an adult has not been shown to improve cross race bias in a multicultural society (Wong et al., 2020). In this study, ethnically diverse Malaysians at a university were tested for facial recognition of the other ethnicities and still showed strong signs of own-race biases with a p-value of less than 0.05 during testing. However, data is currently missing in the young adult and adolescent categories relating to how the environment that one grows up in can affect the solidification of facial recognition skills.

This project seeks to investigate the relationship between people with varying levels of self-reported childhood exposure to different origins, measured by a mean exposure index to different

origins within school, extracurriculars, and their community as and their respective abilities to identify people in a crime of different and same origins. This was compared and studied with the accuracies of identifying the correct individual after watching a video of a crime, simulating a police lineup that occurs if you are a witness of a crime. As such, this would effectively test the application of the reasons behind the cross-race effect in the real world, so that innate factors will better inform courts of potential biases behind eyewitness identification. The expected results are that as exposure to a certain group as a child throughout childhood increases, the ability to recognize faces of that said race will also increase consistent with cross-race bias.

METHODS

Participants were recruited through various methods via a flyer including social media – Instagram account, discord servers, reddit, Nextdoor, and WhatsApp. After filling out a pre-screening form that standardized the cognitive fitness and age of participants, the individuals that identify with an origin group that has video stimuli for testing were asked to fill out the childhood inquiry survey. Previous determinations of cross-race bias often used race as a classification method, however, race is largely a societal construct and should further be specific to geographical region to remain more streamlined with current standards for classifying groups of people. The term “origin” in this study was used to describe a person’s background instead of race or ethnicity and adheres to the United States 1997 Standards for the Classification of Federal Data on Race and Ethnicity with additional subcategories for specificity based off shared

cultural, topographical, and generic diaspora differentiations of larger groups such as “Asians”.

This study used a form of deception to study innate biases. Instead of informing the participants that they are to specifically look for the face, they were informed that this is a study for the purpose of measuring memory in high stress situations. After the conclusion of the study, participants were debriefed, indicating that the study was instead for the purpose of studying witness face recognition of suspects relating to one’s exposure to origins as a child. This extended to the informed consent form and any surveys that the participant takes. A random integer number generator within the bounds of [0,1] was run to select participants should they have filled out the pre-screening form completely and pass the qualifications to participate (see Fig. 1). Due to limitations with permissions on video stimuli, the study was be limited to people of certain origins and was be analyzed accordingly. Through running the random number generator, participants that return the integer number 1 were included in the study, while those with 0 were not.

Those that made it through were contacted to sign forms and given the demographic collection survey on Microsoft forms or Google Forms, depending on participant preference. Although these were given out at the same time, responses were not accepted until an informed consent form was filled out by the proper entities as required. Data collection was set up on the same form after the preceding steps were complete. The pre-screening procedure in the study is displayed in the subsequent graph (Fig 1.)

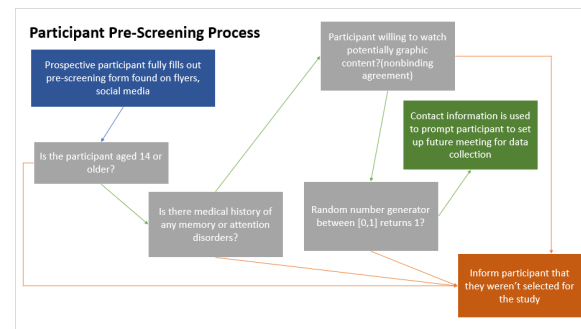


Fig 1. Flowchart demonstrating process between pre-screening form and childhood information survey to collect participants

Collected videos were categorized by the identified race or race stated in the biography of target individual in the video using categories of origin as defined by the previously utilized options for self-reported origin. The participants that have fully signed forms and all previous prerequisites had their responses to Origin and Age recorded for analysis purposes.

Participants took the childhood exposure survey section of the form and were informed about the false purpose of the study. In addition to this, they were also be informed of the approximate length of the study and the content or trigger warnings that may show up within the crime videos. Participants were not asked for their name to maintain anonymity. Origin constitutes designations of the human population into classifications as previously described. The data table of initial childhood survey is shown in Fig 2 and 3. As seen, there were six categories asked about, determined by areas of childhood that experience the most significant amounts of human interaction. These were at school, at extracurriculars, and in the community. Childhood was then further separated into two age groups, 0-9 and 10-18 to better differentiate any nuances in childhood experience

between early childhood and late childhood, but also not too specific so that participants could best remember their exposures. To simplify the rating process, a scale of 0-4 was chosen to best represent a value and correlation in which an individual had very little to no exposure to an origin – 0, and 4 being an origin within the bounds of the specific age group and situation that was defined by “most interactions, if not all, were with this origin”.

Data Table: Selected Sample of Collected Values of Self-Reported Exposure to Primary Origin during the ages of 0-9

Participant #	Self-reported childhood school origin demographic ages 0-9 with an index of 0-4	Self-reported childhood community origin demographic ages 0-9 with an index of 0-4	Self-reported childhood extracurricular origin demographic ages 0-9 with an index of 0-4
22	4	4	4
31	2	2	3
34	2	2	1
43	4	4	4

Fig 2. Data collected during the childhood survey for ages 0-9

Data Table: Selected Sample of Collected Values of Self-Reported Exposure to Primary Origin during the ages of 0-9

Participant #	Self-reported childhood school origin demographic ages 10-18 with an index of 0-4	Self-reported childhood community origin demographic ages 10-18 with an index of 0-4	Self-reported childhood extracurricular origin demographic ages 10-18 with an index of 0-4
22	4	4	4
31	2	2	2
34	2	1	1
43	1	1	1

Fig 3. Data collected during the childhood survey for ages 10-18

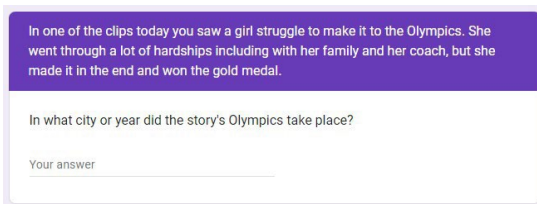
If they cannot remember, they were asked to take their best guess. Each survey question was required to be responded to. To reduce redundancy in the survey, a free response box was provided after rating the primary origin of exposure during one age group to rate on the scale of 0-4 as described above to any other secondary origins that may have had exposure beyond and index of 0. All other origins not mentioned were assumed to have an index of 0,

meaning little to no exposure. The data was taken and an overall index of exposure to an origin was calculated based off an average index of each of the six categories. Each of the categories was weighted equally. An overall index less than 2 was chosen and an overall index greater than 2 was chosen to compare and assigned low exposure index and high exposure index respectively. These categories were chosen and tested for bias by comparing the percentage correct for high exposure bias and low exposure bias. Unlike previous studies, this study did not regard the origin of the participant in its hypothesis; rather, it sought to find a correlation between primary origin of exposure as determined by the highest average index value and its correlation with cross-race bias, rather than correlations between established precedent of own-race bias.

Participants were shown 4 videos of an emotionally significant interaction pulled from creative commons YouTube videos with a clear “suspect” but were not told to pay attention to the suspect, one at a time. The number of videos is significant enough to be differentiable for data analysis but is not overbearing for the participant. 2 were of high exposure, while 2 were of low exposure. The videos utilize the “randomize order” function in respective forms software to prevent primacy and recency effect affecting the memory of individuals. Participants can be asked about any bare, uncovered relevant face from the video not limited to the suspect to serve as a distraction from the intended purpose of the study. Each face was clear for at least 10-60 seconds to be tested, and videos that were tested against each other in the opposing origin categories differed in terms of the

airtime of the face they are asking about by no more than 2 seconds. Faces were front facing and clear in the tested photos, and cropped above the shoulders, preventing the cues beyond the head and face area being used to assist in recognizing the individuals in the stimuli.

Another random generator was used to decide if participants were to have a delay of one day or a few minutes before answering the test questions on the stimuli, representing the variability of when a witness is called into questioning. Participants were asked to schedule a follow-up meeting one day from the scheduled initial meeting to question if they were selected to be part of that group. Each question to be asked on the following day after exposure followed the misleading intent of the study. Two questions were asked regarding each video. Each video had distraction questions that have nothing to do with the intent of the study, for example, “What color was the car?” to make the participant feel like they were taking a general memory test of the stimuli they were exposed to the day before.



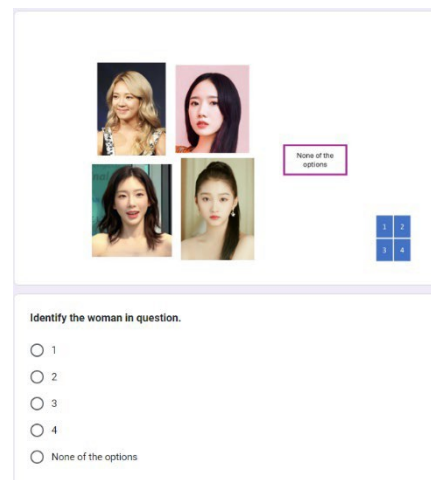
In one of the clips today you saw a girl struggle to make it to the Olympics. She went through a lot of hardships including with her family and her coach, but she made it in the end and won the gold medal.

In what city or year did the story's Olympics take place?

Your answer

Fig 4. Experimental setup of the prompt and distraction question based on the tested stimuli These were a mix of multiple choice and free response to accurately replicate the questioning process. The order of the questions was randomly

selected based on the video stimuli that participants were exposed to. One question, the face identification question, is guaranteed to randomly show up somewhere within those 2 questions. An example of an experimental setup of one of the data questions is shown (fig 4 & 5). The answers for distraction questions were not recorded, nor did they count towards any percentages, rather only to assess the general attention and memory of participants in both the one day and couple minutes delay group. Participants were shown 4 faces relating to each video watched, with only one being the correct option. All selections were standardized by cropping faces above the chest line. Face recognition questions were categorized as either correct or incorrect. Overall, the total percentage of correctly identified faces in each category were calculated and recorded. The overall test was given a time limit of 5 minutes to prevent overthinking.



Identify the woman in question.

☐ 1
☐ 2
☐ 3
☐ 4
☐ None of the options

Fig 5. Experimental setup of face recognition question based on the tested stimuli

DATA ANALYSIS AND RESULTS

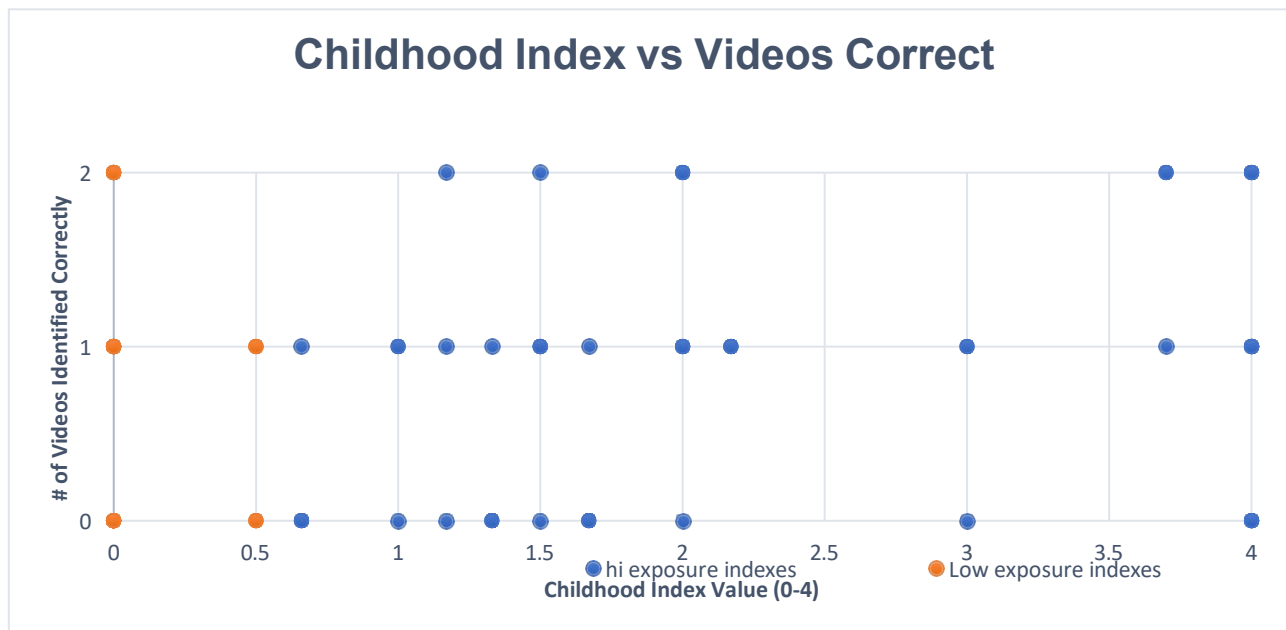


Fig 6. Final graph comparing mean exposure index to number of videos correct

The data analysis process was investigated by way of Pearson's correlation coefficient for investigating a linear relationship, as well as other measures of central tendency. By plotting self-reported numerical exposure against number of same-race suspects correctly identified and number of different-race suspects correctly identified as a percentage, results can be seen if there is any positive correlation with exposure and cross race and same race identifications, shown in the overall graph (Fig. 6). The measurements of the independent variable, against the dependent variable, either same race or different origin identifications was plotted using excel to find an R-value of correlation.

The next step was to analyze a histogram, including categories of increments of [0,1), [1,2), and so on regarding total number of faces identified to see if there any trends in the groups of childhood

index, with the final index being a single category of exclusively 4, which was the group of mostly a single origin's exposure during childhood as reported by the survey. Other tests include isolating individual groups of demographics to see if correlations are stronger in those groups.

Results of the study included overall data from 51 trials from the survey and test. Participants who reported own-origin as two or more origins, unlike other studies, were included in the study since the study investigated childhood exposure rather than relation to own origin. The study demographic included individuals of four different origins, with East Asian being the majority. This includes 18 men and 33 females. The average childhood index of both high and low exposure groups was shown overall in figure 7, with the high number of people with 0-1 indexes due to little to no exposure to the origin tested in low exposure stimuli.

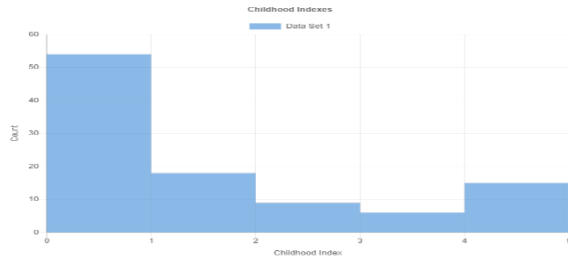


Fig 7. Resulting average indexes for high and low exposure origins collected from participants during childhood survey.

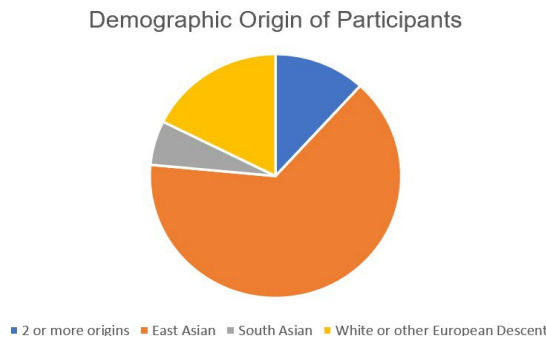


Fig 8. Origin of Participants in the Study

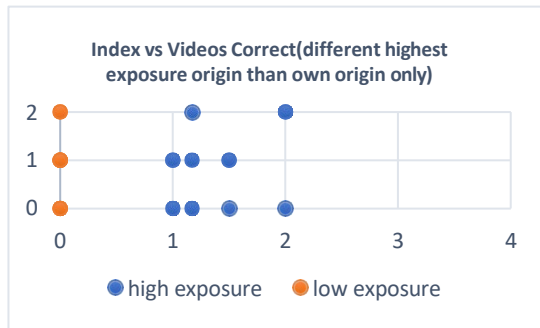


Fig 9. Isolation of data from individuals with a different highest self-reported exposure index to an origin than one's own origin

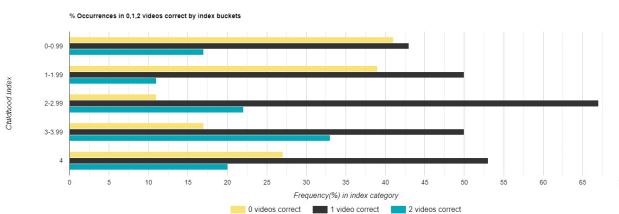


Fig 10. Multiple Frequency Graph of % Occurrences in each Index Bucket

The results of the study showed a slight correlation ($R^2 = 0.19$) between index and

ability to recognize faces when using data and testing from highest and lowest index for origin exposure for the overall data in figure 6, meaning that there was a slight correlation between the two variables (the closer the value is to 1 or -1, the stronger the linear correlation – a value of 0 means no correlation). When only analyzing the correlation values for data points in high exposure, the correlation value was 0.211, meaning there was a stronger correlation found within the high exposure group. However, when observing figure 9, the correlation of values for participants who had a different highest exposure origin than their own origin was higher at 0.293 meaning that the isolation of these values gave a stronger linear correlation than that of the general participant population, showing that childhood exposure index is a stronger predictor of ability to recognize faces, even when not considering participant's own origin and supporting the hypothesis. The average difference between facial recognition of high versus low index facial recognition for any given individual was +0.10, showing favorability towards high childhood index origin, meaning the tests that were of higher exposure faces as indicated by the index were more often identified correct when compared to faces of the lower exposure index origin.

The multiple frequency histogram created compares the percent frequencies of 0,1, and 2 videos correct in each of the index buckets. When comparing the general trend of 0 videos correct, there is a general decrease in the percentage of participants that identify 0 videos correct as index increases. In contrast, the percentage of participants that are able to identify 2 videos correct as index increases also increases. The percentage of

participants who are able to identify 1 video correct as index increases stays relatively the same, if not showing a slight increase.

The overall results show a slight correlation between calculated childhood index and the number of videos identified correctly. Correlation values could have been affected by the lack of sliding scale for dependent variable values. Other methods of analysis showed a better ability to recognize faces from an origin determined to have highest exposure based off childhood index when compared to an origin determined to have lower exposure based off childhood index. In this study, a stronger relationship was shown in the population with a different highest exposure race than one's own origin, highlighting the applicability of the childhood index as a basis for bias in facial recognition. This study also supports previous research on the presence of CRB in individuals ages 14+ as shown through comparing white origin and east Asian origin's abilities to identify faces of white origin. Therefore, the hypothesis that a higher exposure index will result in better facial recognition was only partially supported by the study. Higher exposure as determined by childhood index showed a positive impact on facial recognition. While the correlation between the two variables was not strong linearly, this suggests other factors beyond just childhood(nurture) influences facial recognition abilities, in line with what other previous studies on cross race bias have found.

DISCUSSION

This study has presented a possible root cause to cross-race bias in witness scenarios, specifically investigating childhood exposure and its effect on ability to recognize faces rather than the larger overarching contact hypothesis. While this study has not established precedence in one theory or another, it opens the possibility of other environmental or genetic correlations to cross-race bias, increasing general understanding of what mankind can do to best prevent the implications of it. From an early education perspective, increasing the diversity of classrooms may reduce prejudice in multiracial societies later, although this is not applicable to identifying witnesses.

Presenting this research to courts of law will encourage additional considerations for the reliability of eyewitness evidence in criminal scenarios, pointing to the possibility of flaws in convictions that solely present eyewitness evidence. Future research should also incorporate more origin groups and allow the results to be validated in a more representative sample, as with all research studies. The future studies should also aim to represent a witness scenario better accurately by employing stimuli that is of police bodycam footage, which are real life scenarios. However, these are difficult video to obtain in the United States without obtaining case information. In the future, childhood index could also be calculated using demographic data from previous censuses of countries and neighborhoods to reduce self-report bias and lack of accurate memory

REFERENCES

- Davis, M. M., Hudson, S. M., Ma, D. S., & Correll, J. (2015). Childhood contact predicts hemispheric asymmetry in cross-race face processing. *Psychonomic Bulletin & Review*, 23(3), 824–830. <https://doi.org/10.3758/s13423-015-0972-7>
- Gaias, L. M., Gal, D. E., Abry, T., Taylor, M., & Granger, K. L. (2018). Diversity exposure in preschool: Longitudinal implications for cross-race friendships and racial bias. *Journal of Applied Developmental Psychology*, 59, 5–15. <https://doi.org/10.1016/j.appdev.2018.02.005>
- Hourihaan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2013). Same faces, different labels: Generating the cross-race effect in face memory with social category information. *Memory & Cognition*, 41(7), 1021–1031. <https://doi.org/10.3758/s13421-013-0316-7>
- Innocence Project. (2021). *Explore the Numbers: Innocence Project's Cases and Impact*. Innocence Project. <https://innocenceproject.org/exonerations-data/>
- Otterbourg, K., Stephens, K., Paredes, J., O'Brien, B., & Gross, S. (2022). *Race and Wrongful Convictions in the United States 2022*. National Registry of Exonerations. <https://www.law.umich.edu/special/exoneration/Documents/Race%20Report%20Preview.pdf>
- Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., & Lee, K. (2014). On the Links Among Face Processing, Language Processing, and Narrowing During Development. *Child Development Perspectives*, 8(2), 65–70. <https://doi.org/10.1111/cdep.12064>
- Trawiński, T., Aslanian, A., & Cheung, O. S. (2021). The effect of implicit racial bias on recognition of other-race faces. *Cognitive Research: Principles and Implications*, 6(1). <https://doi.org/10.1186/s41235-021-00337-7>
- Wong, H. K., Stephen, I. D., & Keeble, D. R. T. (2020). The own-race bias for face recognition in a multiracial society. *Frontiers in Psychology*, 11(208). <https://doi.org/10.3389/fpsyg.2020.00208>
- Wright, D. B., Boyd, C. E., & Tredoux, C. G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology*, 17(3), 365–373. <https://doi.org/10.1002/acp.898>

Interplay of Individual Personality and Language

Background on Speech Accommodation

Kamden Applebee

ABSTRACT

When two individuals of different native languages communicate, there are accommodative behaviors (such as choices in vocabulary, slightly slower speech, and nonverbal, clarifying gestures) practiced by both speakers to create understanding and work towards conversational goals. This correlational study seeks to examine the relationship between an individual's personality, their attitudes about performing these accommodative behaviors, and their level of bilingualism. A survey was constructed referencing three existing questionnaires to create an assessment of an individual's self-reported accommodation attitudes, measure of their bilingualism, and their personality based on the Ocean Big 5. A sampling procedure was conducted, and participants (n=61) completed the survey. This data was processed using the correlational software RStudio, and found multiple statistically significant r-squared values, including a moderately strong (-0.41) inverse correlation for accommodation and bilingualism, which suggests individuals who consider themselves the most speech-accommodative are also the least bilingual. A considerable (-0.61) direct correlation was also observed between speech accommodation and extraversion, suggesting individuals who consider themselves extroverted and try frequently to communicate with others do not feel as accommodated. The hypothesis of a connection between extraversion and accommodation, as well as accommodation and bilingualism, has been proven correct. In the future, the study could ideally be conducted again with randomizations to the question order within each section, and with a larger range of participants, to eventually produce a greater understanding of the connection between personality, speech accommodation, and bilingualism.

INTRODUCTION

This correlational study seeks to examine a connection between an individual's OCEAN personality types, their speech accommodation attitudes, and their level of bilingualism. It would prove valuable to showcase a positive correlation between any two of these three factors, because proving a connection will allow for stronger evidence of the benefits to teaching multiple languages. For example, if there is observed to be more speech accommodation done by people who are more bilingual, there would be evidence to support that early introduction to other languages

can lead to better speech accommodation of others. Or, if more bilingualisms has a significant relationship with a specific set of OCEAN personality types, it could be suggested that there are benefits to an individual's overall identity by the teaching of additional languages and skill of bilingualism.

There are a multitude of factors that contribute to racism, both historically within specific societies at large, and psychologically within certain individuals. The impact of these factors persists in many people's day-to-day interactions, despite the fact our world contains more than 600 diverse

ethnicities -and an estimated 43% of people in the entire world are bilingual (Elizabeth, 2023), providing a clear basis for the necessities of inclusion and accommodation towards others. One notable psychological influence on an individual's racial attitudes is outgroup bias, the unconscious human tendency to assume that all groups outside your own share certain characteristics, which can predispose your thinking of others to revolve around stereotypes and shallow generalizations. You may, for example, think that all individuals who are not native to English are less intelligent, or need to be spoken to like children in order to be understood. Whether you hold these biases or not, when in conversation with individuals who are not native to your own language group, you may alter (or consider altering) your speech for the sake of being perceived with more clarity, and to avoid possible frustrations and confusions. These speech-accommodating adjustments tend to occur from both speakers of a conversation, with varying success towards facilitating ease of communication. Interestingly, as recent research has shown, individual personality and language backgrounds can influence these accommodative behaviors. Individuals who are even slightly bilingual have been shown to be better at speech accommodation (Kathrin Rothermich, et al., 2023), because people who have been accommodated themselves are more precise with their own accommodation: they know how simple is too simple, without being condescending towards the individual being accommodated. However, we have limited data on the interplay of these factors, such as the extent of correlation between speech accommodation attitudes and individual personality.

One study that previously examined the cross-cultural relationships of personality traits was done in 2006, by Antonia Terracciano and Robert R. McCrae. In this study, measures of the big-five personality spheres were examined across various geographical locations, yielding that no particular personality type was more prevalent based on location or language spoken alone. This further suggests that human personality types (as measured by the five-factor model of personality, FFMP) are equally common across varying cultures. The study showed also showed differences in FFMP across sexes, though, with men scoring higher only by about half a standard-deviation on assertiveness, excitement-seeking, and openness to ideas, while women scored higher on warmth, gregariousness, and openness to aesthetics and feelings. These tendencies suggest that although traits are organized similarly across cultures (Williams & Best, 1990), there are pancultural patterns associated with age and gender that tend to largely correlate with North-American literature and gender stereotypes, as opposed to the specific language an individual was raised to speak.

METHODOLOGY

To create a survey suitable for examining this project's three facets, multiple existing surveys were referenced and modified. The structure of the survey begins with simple demographic questions, then continues to examine the participant's specific language background, then ends with a segment on individual personality.

Figure 1: Survey Segment

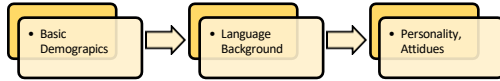


Figure 1: Graphic showing the structure of the survey's main three segments.

The first of these surveys is the Language and Social Background Questionnaire - Revised (LSBQ-R) (Anderson et al., 2018), which was altered from its previous iteration to assess limitations of the previous Language and Social Background Questionnaire (Luk & Bialystok, 2013). The LSBQ-R is composed of 22 questions in three sections. First, it aims to gather basic demographic information (such as age, sex, and race), and secondly, examines which language(s) the participant speaks and/or can understand; and third, to assess language use in different stages of life and atmospheres. Within this study, 12 questions from the LSBQ-R are utilized to collect basic information about the participant's age and ethnicity, and then, to collect metrics regarding the participant's bilingualism. They are asked to rate their bilingualism on a scale of one to 100 (one being monolingual and 100 being bilingual), and to rate on a scale of 1 to 5 (one being only English and five being never English) how often they speak English versus another language(s) at home, at school, and at their place of work. This range of settings, as sourced from the original LSBQ-H, aims to explore fluctuating stances regarding different social spheres the participant is likely to use language within.

The second survey referenced in this study is the Speech Accommodation Questionnaire (SAQ) (Rothermich et al., 2023). The SAQ, designed for a

study titled “Reported Use of Second-Language Speech Accommodation in Everyday Interactions: The Role of Individual Differences”, is an online questionnaire that utilizes two distinct wordings: asking about the accommodative behaviors of the participant themselves, compared to asking about the accommodative behaviors of people in general. By comparing the participant's outlook on themselves versus their outlook on others, the study can assess and anticipate differences between inner ideals and actual actions. Within this study, the questions have been modified with two distinct wordings: with the perspective directed at L1 English speakers, and separately, directed at L2 English speakers. As the structure in [Figure 1](#) shows, this study's use of the SAQ seeks data regarding two separate perspectives, in order to compare the self-perceived attitudes of more proficient speakers with the actual perceptions of less proficient speakers. As shown in [Figure 2](#), the original SAQ (left column) examines the difference in speech accommodation attitudes when a participant answers questions about themselves, written in first person; versus questions about people in general, written in third person. The modified SAQ (right column) examines the difference in individual perspectives of native speakers and of non-native speakers. When participants first enter the study, following their submission of the informed consent form, they are then prompted to select between the survey for native English speakers, and the survey for non-native English speakers. This way, all volunteers receive the correct question set, and all responses are completely detached (and made anonymous) by the name submitted to the informed consent document.

Figure 2: Example of Modifications to Original SAQ

Q1 - 1st When I talk to a less proficient English speaker, I adjust my speech to improve my chances of successfully conveying a message. ☐ 1- Strongly Disagree (1) ☐ 2 - Disagree (2) ☐ 3 - Neutral (3) ☐ 4 - Agree (4) ☐ 5 - Strongly Agree (5)	When I talk to a less proficient English speaker, I adjust my speech to improve my chances of successfully conveying a message. * 1 2 3 4 5 Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree
Q1- 3rd When people in general talk to a less proficient English speaker, they adjust their speech to improve their chances of successfully conveying a message. ☐ 1 - Strongly Disagree (1) ☐ 2 - Disagree (2) ☐ 3 - Neutral (3) ☐ 4 - Agree (4) ☐ 5 - Strongly Agree (5)	When a more proficient English speaker talks to me, they adjust their speech to improve chances of successfully conveying a message. * 1 2 3 4 5 Strongly disagree ☐ ☐ ☐ ☐ ☐ Strongly agree

Figure 2: An example of the modifications made to existing Speech Accommodation Questionnaire questions for the purpose of this investigation. Examples on the left are from the study that originated the SAQ, and examples on the right are from this study.

The third survey utilized is the Ten Item Personality Index (TIPI) (Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr., 2003), which is designed to encapsulate the participant’s ratings of their own social identity based on the five factors personality dimensions: extroversion, agreeableness, conscientiousness, emotional stability, and openness (McCrae and Costa, 1989). The participant answers ten questions, with each of the five dimensions relating to two questions. One of these questions has its value (from one to seven) inverted and is averaged with the unchanged response from the other question, to generate 5 values each on a seven-point scale that quickly and numerically describe a participant’s personality.

Using these sets of questions, a survey was constructed containing three segments, as shown in [Figure 1](#). Flyers for this survey were posted around a

high school, as well as distributed to online groups containing members living in both the United States and United Kingdom, and to college students in Germany for the sake of broadening the demographics sampled. Before accessing the actual survey materials, all participants were prompted to read and accept a voluntary consent form and provide their digital signature (or their legal guardians, if the participant indicated they were under 18). Then, from this initial Google Form, they were directed to a form for speakers native to English, or to a form for learners of English. The 1st and 3rd sections of these surveys were identical, while the 2nd section diverged as sampled in [Figure 2](#). This pair of surveys collected no personally identifying information, and all responses were shuffled by a random number generator for analysis. In total, (n=61) participants submitted responses,

with (n=41) to the native-English form and (n=18) to the non-native-English form.

Figure 3: Example of Data Table

Survey From:	RL1
Section 1	Basic Demographics
Biological Sex	=Form Responses 11B2
Years Old	=Form Responses 11C2
Racial/Ethnic Heritage	=Form Responses 11D2
Language(s) in home	=Form Responses 11E2
First language acquired	=Form Responses 11F2
Other langs learned	=Form Responses 11G2
Dominant lang past 3yr	=Form Responses 11H2
Most often 2nd lang	=Form Responses 11I2
Use at home 1eng-5oth	=Form Responses 11J2
Use at school 1eng-5oth	=Form Responses 11K2
Use at work 1eng-5oth	=Form Responses 11L2
Bilingualism 1rone-100	1
Bilingualism Score 0.03 to 1	=(SUM(B14:B16)+B17)/115 =ROUND(B18,2)
Section 2	Speech Accommodation Att
1they adjust to improve	=Form Responses 11N2
1they fac. understanding	=Form Responses 11O2
1they conversation goals	=Form Responses 11P2
1they reduce conflict	=Form Responses 11Q2
1they avoid judgement	=Form Responses 11R2
1they arrogant tone	=Form Responses 11S2
1they talk down	=Form Responses 11T2
1they avoid adjusting	=Form Responses 11U2
SAA Score from 1 to -1	=(SUM(B21:B24)-SUM(B25:B28))/11 =ROUND(B29,2)
Section 3	Personality TIPI
Extraverted, enthusiastic	=Form Responses 11V2
Critical, quarrelsome	=Form Responses 11W2
Dependable, self-discipline	=Form Responses 11X2
Anxious, easily upset	=Form Responses 11Y2
Complex, open to new exp	=Form Responses 11Z2
Quiet, reserved	=Form Responses 11AA2
Warm, sympathetic	=Form Responses 11AB2
TIPI value (cumulative dist)	

Figure 3: After received in either of the Google Forms (English as a primary, L1 speakers, or English not as a primary, L2 speakers), order of responses is shuffled by a random number generator, and each value is sorted into the per-section data table. This Google Sheet calculates values for the participant's Bilingualism Score (from 0.03 to 1), Speech Accommodation Score (from -1 to 1), and cumulative TIPI value (from X to X).

Using these responses in a Google Sheet, each of the seven values (one from each of the first two sections, then five from the TIPI) are calculated and used to construct a correlational graph. In total, there are three graphs: one comparing Bilingualism and SAA, one comparing Bilingualism and

cumulative TIPI, and one comparing cumulative TIPI and SAA.

RESULTS

The hypothesis of a measurable relationship between some aspects of personality, speech accommodation, and bilingualism, has been proven partially correct.

Notable correlations within the overall sample ($n=61$) include a moderately strong inverse correlation (-0.46) between accommodation and bilingualism, indicating that participants who found themselves less bilingual also found themselves more accommodative. This is likely because the majority of the volunteering participants were L1 (native) speakers of English, who are likely to respond to a survey in English without necessarily being bilingual at all.

Figure 4: Correlation Matrix of All Participants

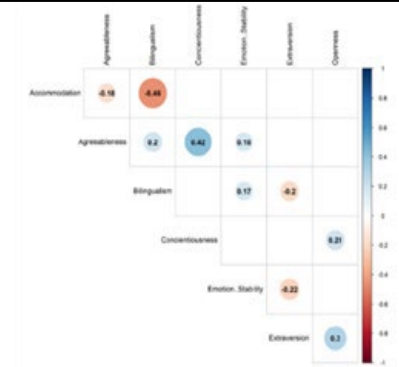


Figure 4: A correlation matrix constructed in R (RStudio) depicting 21 comparisons between all variables available across all participants available

Comparing participants who were male and female, the moderately strong inverse correlations (-0.42 on male, -0.45 on female) between Accommodation and Bilingualism persisted.

The strongest correlation in the female sample was a strong direct correlation (0.55) between Agreeableness and Conscientiousness, while the strongest correlations in the male sample included Bilingualism and Openness (-0.6), suggesting males

in this sample who are less bilingual also consider themselves more open, while females in this sample did not have any notable correlation between Bilingualism and Openness

Figure 5: Correlation Matrix of Male vs Female Participants

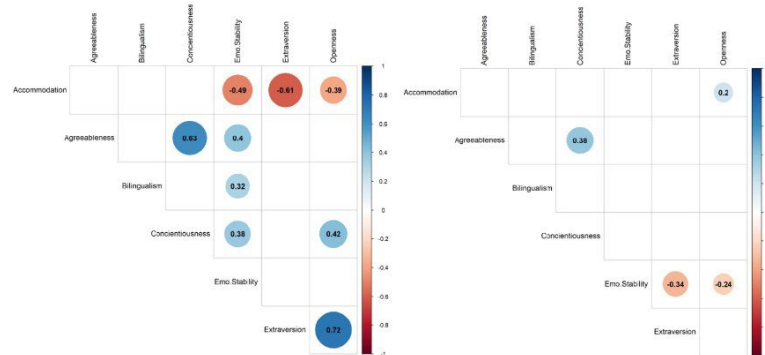


Figure 5: A correlation matrix constructed in R (RStudio) depicting 21 comparisons between all variables available, the set on the left from all male participants (n=22) and the set on the right from all female participants (n=36).

There were numerous of strong correlations within the L2 sample, likely because of its small size (n=18) compared to the L1 sample (n=43). Notably within the L2 sample is a strong inverse correlation (-0.61) between Accommodation and Extraversion, suggesting that L2 individuals who feel they are more extraverted and try more often to interact with other people also feel that they are accommodated more poorly by those people. This further implies that L2 individuals who choose to be extraverted in a language that isn't native to them are often poorly received by others, and that communication can be made difficult by those they try to interact and create connections with. L2 individuals who felt they were poorly accommodated also tended to feel more emotionally stable, with a moderately strong inverse correlation of (-0.49). This, combined with the

strong inverse relationship between Extraversion and Accommodation (-0.61) suggests that the poor speech accommodation experienced by L2 individuals was frequent with individuals who felt they were both extraverted and emotionally stable. It additionally suggests that these individuals who felt poorly accommodated likely did not feel that way as a result of emotional oversensitivity or a lack of attempts to communicate socially with others.

Figure 6: Correlation Matrix of L2 vs L1 Participants

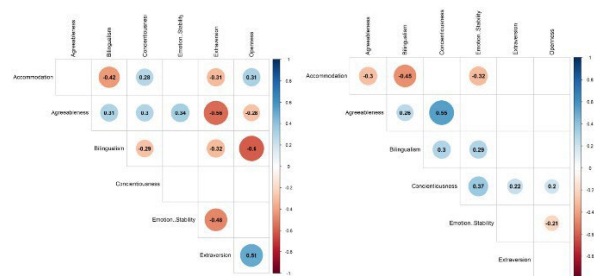


Figure 6: A correlation matrix constructed in R (RStudio) depicting 21 comparisons between all variables available, the set on the left from all L2 participants (n=18) and the set on the right from all L1 participants (n=43).

In the future, this study plans to reconduct its procedures with a larger sample and better balance between L1 and L2 speakers, as well as male and female participants. To broaden the hypothesis, expansion down the line could perhaps use other personality tests to create its variable on individual personality, such as tests related to neurological disorders or the structure of a participant's family. There is also a potential to expand the focus by

examining regional location, age, and additional quantifications of its variables, such as course credit from taking second-language courses as a measure of bilingualism. If it is reputably argued that there are benefits from bilingualism (such as positive personality traits, and more accommodative behaviors towards others), it could motivate earlier introduction of second-language classes during the critical period for language acquisition (between zero and seven years old) for the betterment of broader society; in the form of more acceptance and inclusion towards others.

BIBLIOGRAPHY

- Anderson, John et al. "The Language and Social Background Questionnaire: Assessing Degree of Bilingualism in a Diverse Population. Supplementary Materials". 25 Oct. 2018. Web. 2 Nov. 2023.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. *Journal of Research in Personality*, 37, 504-528.
- Gratton, Elizabeth. "Bilingualism Statistics in 2022: US, UK & Global." Preply, 13 Oct. 2023, preply.com/en/blog/bilingualism-statistics/.
- Rothermich, Kathrin, et al. "Reported use of second-language speech accommodation in everyday interactions: The role of individual differences." *Journal of Speech, Language, and Hearing Research*, vol. 66, no. 9, 2023, pp. 3307–3327, https://doi.org/10.1044/2023_jslhr-23-00167
- Terracciano, Antonio, and Robert R McCrae. "Cross-cultural studies of personality traits and their relevance to psychiatry." *Epidemiologia e psichiatria sociale* vol. 15,3 (2006): 176-84. doi:10.1017/s1121189x00004425

Is it Actually That Dang Phone?

The Effect of Social Media Usage on Emotional Response to Art

Marilyn Zhao

ABSTRACT

Social media usage is often blamed as one of the main culprits of our collective lack of attention during anything from spending time with loved ones to listening to a lecture. But how much of a role does social media usage play in our ability to process art? This project seeks to investigate the effect of social media usage on emotional response to visual art. Individuals use social media to fill time and cure boredom, leading to the loss of time used to process emotions and a decreased attention span. The ability to process emotions as well as sustained attention are vital components of art appreciation, making the potential impact of social media usage on our ability to understand art deeply concerning. Participants (n=140) were asked to observe five artworks before completing a survey utilizing the Multidimensional Emotional Questionnaire (MEQ) to assess emotional response. Social media usage was self-reported and defined as the usage of apps designed for digital social interaction which also possess the infinite scroll feature. Preliminary statistical analysis revealed a trend of positive emotional response decreasing as social media usage group increased and little to no variation in negative emotional response between all usage groups. Further analysis examining artwork-specific differences in emotional response measures (frequency, intensity, persistence, regulation) between the three groups revealed consistent patterns of emotional response between usage groups by artwork, but no fully consistent difference between usage groups across artworks.

INTRODUCTION

This project seeks to investigate the effect of social media usage on emotional intensity in response to visual art through an application of the Multidimensional Emotional Questionnaire (MEQ). By understanding potential differences in emotional response, there can be valuable conclusions made about the overall reception and acceptance of art related to varying levels of social media usage. This knowledge will allow individuals ranging from hobbyist artists to professional marketing teams to better understand the state of their audiences' emotional reactivity.

With addictive entertainment, endless information, and constantly shifting innovations, social media's perks have drawn in 5.17 billion people worldwide since its invention (Shewale, 2024). With a staggering 63.82% of the global population utilizing social media, this method of communication has a strong influence on many lives (Shewale, 2024). However, the full extent of negative impacts of social media are still unknown, especially in relation to the art industry.

An essential skill needed to analyze art—sustained attention span—has been declining with the rising usage of social media and digital

tools. A 2015 Microsoft survey revealed a concerning decrease of attention span from twelve seconds in 2000 to eight seconds in 2013, lower than the attention span of a common goldfish (Gausby, 2015). Additionally, social media is used as a convenient distraction from reality. With social media's constant stream of intense news and information desensitizing consumers to current events, social media is also used to escape emotion and fill boredom. This way of utilizing social media has been shown to lead to concerning impacts on mental health. Research conducted by the Vita Health Group in the UK linked scrolling before bed to decreased quality of mental health. The study additionally found that a concerning 76% of individuals aged from 18 to 24 years old surveyed do not put aside time to process their emotions, citing scrolling on social media as a primary reason why (Bivins, 2021). Both of these factors – decreased attention span and decreased time spent processing emotion – may translate to a decreased ability for individuals with high social media usage to process art fully and emotionally. As the definition of art is most often centered around a piece's ability to provoke emotion, this can be understood as a decreased appreciation for art.

Although there are studies mentioned above have investigated the effect of social media/screen time on mental processes that will likely affect the perception of art, there has been little research directly unpacking the potential effect of social media usage and its consequences on emotional response to art stimuli.

A decreased emotional response to art is deeply concerning across many industries and fields: traditional visual artists will need to work even harder to live off of their work, marketing will be forced to shift as advertising is a form of artwork, art-informed history will suffer, and all gaming and television will eventually change to fit this new numb norm. With 4.43 million individuals in the U.S. alone employed in the arts as of 2022, there are numerous individuals worldwide at risk of having their very livelihoods compromised by this shift (Statista, 2022).

As social media addiction and usage are on the rise, public concerns surrounding the negative impact of social media on appreciation of traditional arts—especially in relation to the recent popularity of AI art—have also been increasing. Art as an industry is constantly shifting and deeply impacted by economic factors: corporate demand for art expertise changes and consumer's financial ability to invest in a luxury good varies. In our high stimulation technological age, the concerns around an overall lost appreciation for the arts have grown from a mild unease to genuine fear. Having less emotional responses to art may signify either a disinterest in artwork or an inability to feel an emotional response, two factors that will deeply affect the industry as social media usage inevitably continues to affect people's lives for years to come.

In addition to consumers being less receptive to art affecting the individual consumer themselves, artists will also suffer if this numbing

effect continues. Artists have historically struggled with financial support in the workplace. A survey on artists in the UK and Ireland performed by a-n, an information company focused on art, found that the average annual income of artists was about £10,000— less than a third of the average annual income (Doeser et al. 2014). Artists surveyed in the UK reported an average hourly wage of £2.60, a figure dramatically below the UK’s set minimum wage of £9.50 (Lola et al., 2023). Additionally, DHA Communications surveyed 1,061 artists and found socially engaged practice as one of the most important factors to earning income (DHA Communications, 2013). What these statistics tell us is that social media is not just a side hobby for artists, but in many cases a highly valuable and possibly essential tool to their career due to its financial accessibility and social popularity. If social media users are becoming less impacted or emotionally invested in the arts, the consequences of that shift for artists struggling to reach a platform online to support themselves will only grow.

To counteract this concerning change, the art industry must first understand the full scope of the issue at hand. Understanding current research, we hypothesize that those with greater social media usage will respond with less emotional intensity to art stimuli than those with less social media usage because of the emotionally numbing effect of excessive social media usage.

BODY— METHODOLOGY

Sampling Procedure

A flyer including a link to the survey form was first distributed to various online and physical locations, including community bulletin boards, school hallways, e-mails, and promotion through social media. The form included a section for reporting screen time and social media usage time with clear instructions for both Apple and Android users, as well as the artworks and MEQ itself.



Figure 1. The flyer used in this study included a QR code with a direct link to the survey. The flyer was posted online as well as printed and posted physically.

Participants who filled out the interest form were divided into groups of low-social media users (below the American average of 127 min), above average social media users (usage above 127 min but below 150% above 127 min/317.5 minutes), and significantly above average social media users (usage above the 150% above average mark/317.5 minutes). The American daily average used to distinguish these groups comes from a 2023 study by Forbes Advisor. Usage being 150% above average translates into about one-third of waking hours

being spent on social media. Participants were given instructions within the form to ensure that initial data collection concerning social media usage was accurate. Names or other identifying

information was not collected to preserve confidentiality and results were numbered, and informed consent forms were embedded

Table 1. *Average screen time and social-media utilization data by usage group.*

	Average daily non-work screen time usage (minutes)	Average daily social media usage over past recorded week (minutes)
Below average usage (< 127 minutes)	224.1515	57.8182
Above average usage (between 127 and 317.5 minutes)	361.5849	213.3145
Significantly above average usage (>317.5 minutes)	584.0000	549.7143

Art Stimuli Selection

Traditional western visual artworks were selected from the art database Artstor to represent art typical of American museums. Digital artworks were randomly selected from Unsplash—another art database—and added to the stimuli pool in consideration of the type of art

most social media users would be seeing on social media apps. The categories are as follows:

1. Impressionist art
2. Post-impressionist art
3. Abstract art
4. Surrealist art
5. Digital art



Figure 2. *The final five chosen artworks from impressionist, post-impressionist, abstract, surrealist, and digital art categories. Numbering corresponds both to the art style category they were chosen from as well as the order they appeared within the survey. Artworks were not numbered in the survey itself.*

Artworks were randomly selected using a random number generator under the search matching the category names above. Ten artworks per category were given to a qualified individual with a Bachelor of Arts in art history to filter. Five artworks per category were selected to be shown to participants as the most representative artworks of their classification (ex: the five most Impressionist artworks). Inaccurately labelled artworks that did not fit the art style category they were digitally tagged as were also filtered out in this stage. Artworks that could not be shown to participants were removed before one artwork was randomly selected from the five representative artworks. Only one artwork was removed in this way, as it displayed nudity.

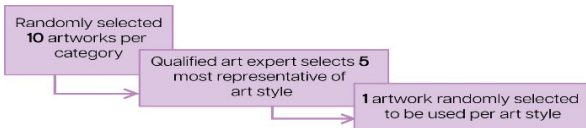


Figure 3. A flowchart explaining the process taken to select and filter artworks. 10 artworks were first randomly selected from sites with free-to-use licenses such as Artstor and Unsplash, filtered by a qualified art expert, and then finally randomly selected again for one final artwork per category.

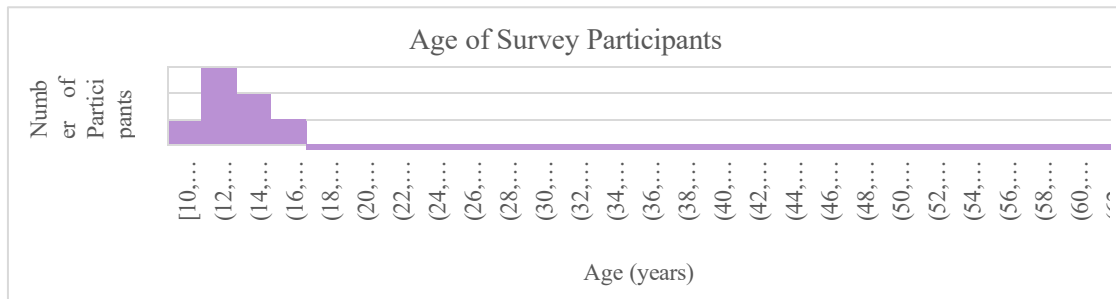


Figure 4. Distribution of survey participant's age. The average age was 16.979 years old, and the median age was 15 years old.

For the below average usage group, the average age was 17.045 years, with the ratio of screen time to social media usage being 3.877. For

PARTICIPANTS

The average age of participants ($n=140$) was 16.979 years old, and the median age was 15 years old. Of the individuals who submitted the interest form, 66 were classified as below average media users, 53 were classified as above average social media users, and 21 were classified as significantly above average media users. 161 participants initially responded to the survey and 21 were removed from inaccurately reporting daily screen times (ex: 28 hours a day). Participants who reported high levels of social media usage (ex: 20 hours a day) were not removed, as we cannot assume that number was misreported based off of the limited information collected and the inclusion of apps like YouTube that may not require active usage in the definition of social media.

the above average usage group, the average age was 17.660 years old, with the ratio of screen time to social media usage being 1.695. For the

significantly above average usage group, the average age was 15.048, with the ratio of screen time to social media usage being 1.062.

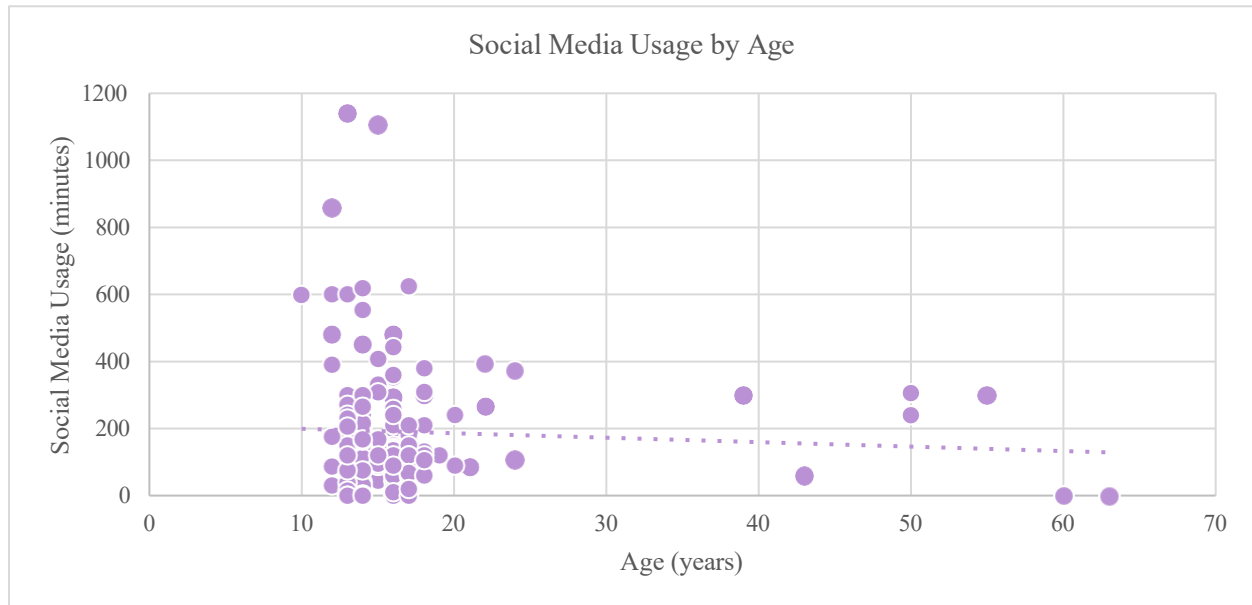


Figure 5. Distribution of social media usage by age. The ratio of screen time to social media usage decreased with each social media usage group (below average = 3.877; above average = 1.695; significantly above average = 1.062).

EXPERIMENTAL PROCEDURE

The Multidimensional Emotion Questionnaire (MEQ) was used to measure emotional response in this survey (Klonsky et al., 2019). This index was administered digitally through Google Form’s Likert scale question format. Time scales were adjusted for the “How Often?” and “How Long-Lasting?” categories from weeks and months to number of times the emotion was felt and seconds/minutes using the same principles of specific quantified numbers originally used in the construction of the MEQ, as seen in Figure 1. The MEQ measures frequency, intensity, persistence, and regulation for 10 discrete emotions (5 positive: happy, excited, enthusiastic, proud, inspired; 5 negative: sad, afraid, angry,

ashamed, anxious). MEQ measures are calculated by summing responses—for example, positive frequency would be calculated by summing all responses to the question “How Often?” for positive emotions. The MEQ also provides the basis for summing measures of frequency, intensity, and persistence to get a measure of overall positive/negative emotional response.

Participants were given an unregulated amount of time to process and view artworks before moving on to fill out the MEQ for each of the five artworks. The instructions given to participants are as follows: “You will be asked to observe an artwork for as long as you feel needed before continuing to answer questions about your

emotional response to said artwork. Here is an example of the questions to clarify how the questioning system works:

	1	2	3	4	5
How Often?	Never	Once or Twice	Often	Nearly Constantly	Constantly
How Intense?	Very Low	Low	Moderate	High	Very High
How Long-Lasting?	Less than 1 second	1-10 seconds	1 minute	1-3 minutes	Longer than 3 minutes
How Easy to Regulate?	Very Easy	Easy	Moderate	Difficult	Very Difficult

This response would mean that you felt decently happy in response to the artwork almost constantly, which persisted for quite a while but was easy to keep in check. There are 5 artworks total.”

The final survey design included informed consent, self-reported screen time and social media statistics, the artworks and MEQ itself, as well as a question asking if participants recognized any of the artworks from previous experiences. None of the artworks stood out in being recognized by many participants. 15 participants reported some feeling of familiarity towards the artworks, but few recognized the specific artworks.

Initial analysis consisted of first grouping participants by their usage group, averaging their initial responses, and then calculating MEQ

measures (frequency, intensity, persistence, regulation) for positive and negative emotions following traditional MEQ guidelines using said averages. This data was then graphed for every emotional measure (including the measure of overall which is a sum of frequency, intensity, and persistence values).

	1	2	3	4	5
How Often?	None	Once or Twice	Often	Multiple Times Throughout	Nearly Constantly
How Intense?	Very Low	Low	Moderate	High	Very High
How Long-Lasting?	Less than 1 second	1-10 seconds	1 minute	1-3 minutes	Longer than 3 minutes
How Easy to Regulate?	Very Easy	Easy	Moderate	Difficult	Very Difficult

Figure 6. Altered MEQ key. The options for “How Often?” and “How Long-Lasting?” were adjusted to fit the purposes of the study.

- Below Average
- Above Average
- Significantly Above Average

Figure 7. The colors utilized to represent different social media usage groups. Groupings: below average = below 127.5 minutes daily (n=66); above average = between 127 minutes and 317.5 minutes daily (n=53); significantly above average = above 317.5 minutes daily (n=21).

Table 2. Average overall positive emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
1	32.63636	30.32075	33.85714
2	24.04545	23.98113	23.2381
3	29.0303	29.66038	27.42857
4	23.74242	23.98113	20.52381
5	31.16667	27.90566	28.42857

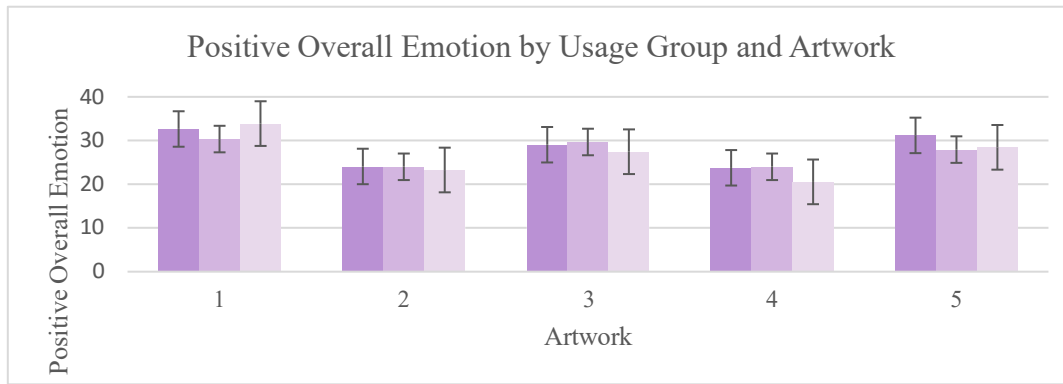


Figure 8. Average overall positive emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 3. Average overall negative emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
1	19.36364	19.13208	22.33333
2	24.48485	23.43396	24.47619
3	20.86364	19.67925	20.09524
4	23.84848	23.96226	24.33333
5	25.12121	24.43396	22.09524

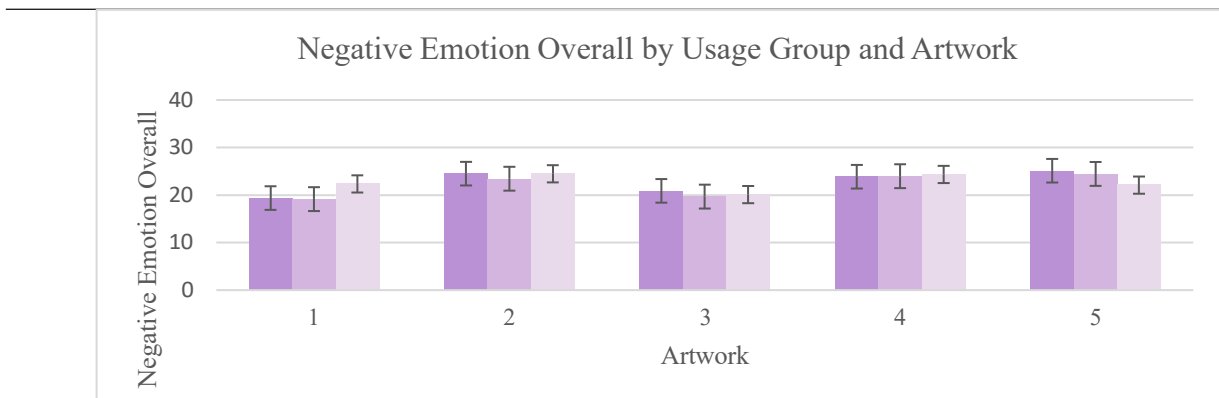


Figure 9. Average overall negative emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 4. Average frequency of positive emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
1	11.43939	11.0566	12.14286
2	8.287879	8.584906	7.809524
3	10.21212	10.73585	9.52381
4	8.333333	8.698113	6.952381
5	10.84848	10.4717	9.952381

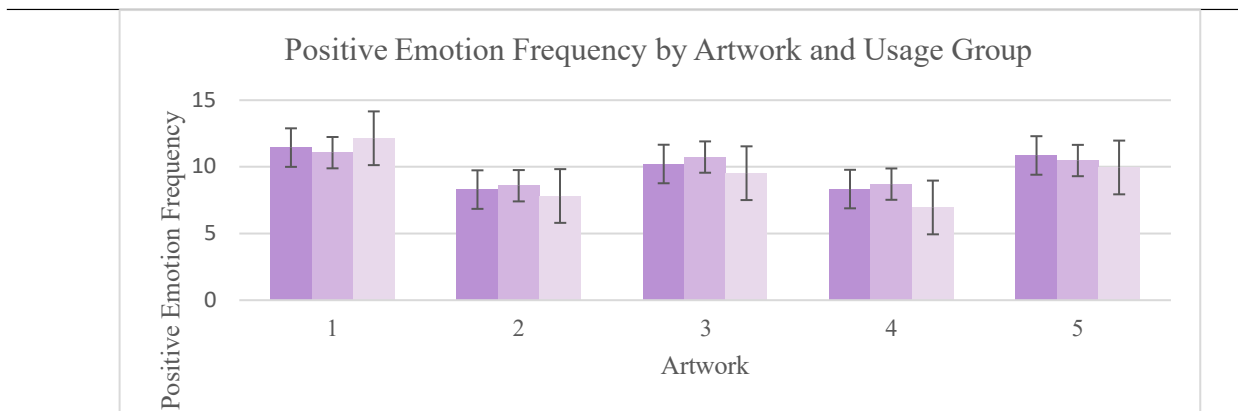


Figure 10. Average frequency of positive emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 5. Average frequency of negative emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
1	6.590909	6.566038	7.238095
2	8.515152	8.264151	8.380952
3	7.181818	6.924528	6.809524
4	8.469697	8.867925	8.714286
5	8.924242	8.90566	7.333333

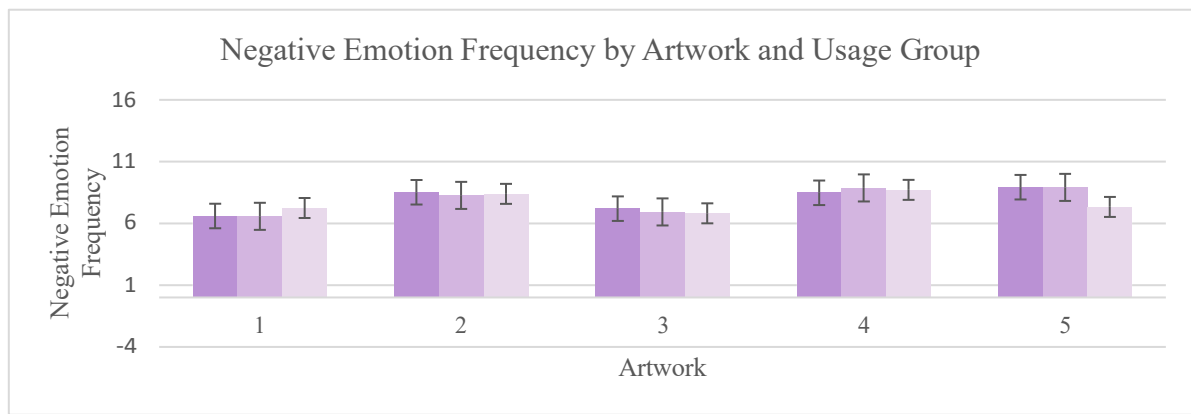


Figure 11. Average frequency of negative emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 6. Average intensity of positive emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
		average	
1	10.56061	9.981132	11
2	7.939394	7.867925	7.714286
3	9.636364	9.754717	9.142857
4	7.636364	7.792453	6.904762
5	10.33333	9.018868	9.428571

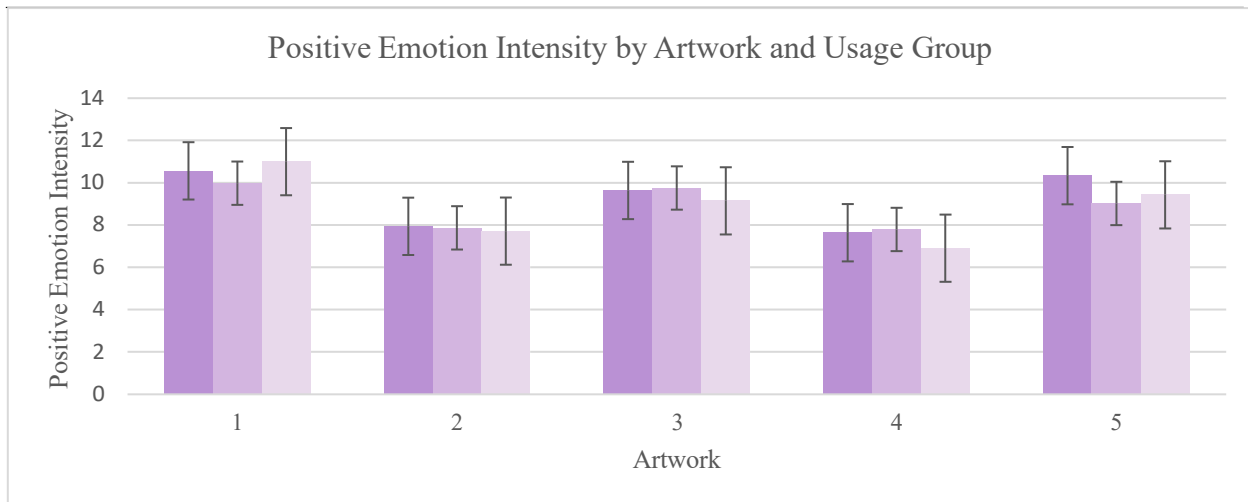


Figure 12. Average intensity of positive emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 7. Average intensity of negative emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
1	6.393939	6.226415	7.666667
2	8.075758	7.603774	8.238095
3	6.833333	6.490566	6.809524
4	7.818182	7.830189	7.571429
5	8.151515	7.962264	7.714286

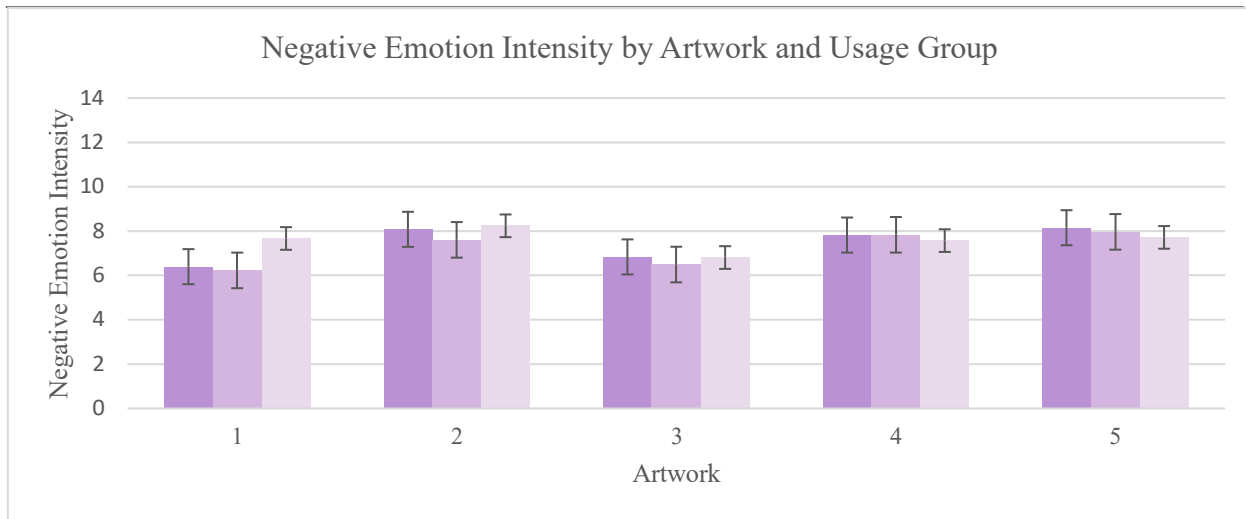


Figure 13. Average intensity of negative emotion response by usage group and artwork. N=140 participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily (n=66); above average = between 127 minutes and 317.5 minutes daily (n=53); significantly above average = above 317.5 minutes daily (n=21). Emotional response assessed by the MEQ.

Table 8. Average persistence of positive emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
1	10.63636	9.283019	10.71429
2	7.818182	7.528302	7.714286
3	9.181818	9.169811	8.761905
4	7.772727	7.490566	6.666667
5	9.984848	8.415094	9.047619

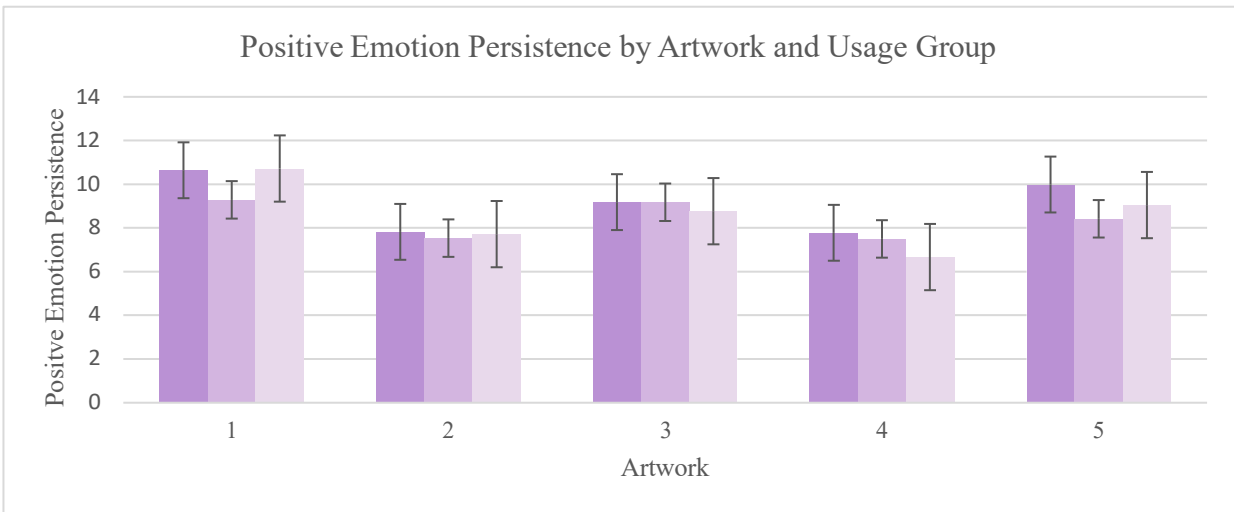


Figure 14. Average persistence of positive emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 9. Average persistence of negative emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
		average	
1	6.378788	6.339623	7.428571
2	7.893939	7.566038	7.857143
3	6.848485	6.264151	6.47619
4	7.560606	7.264151	8.047619
5	8.045455	7.566038	7.047619

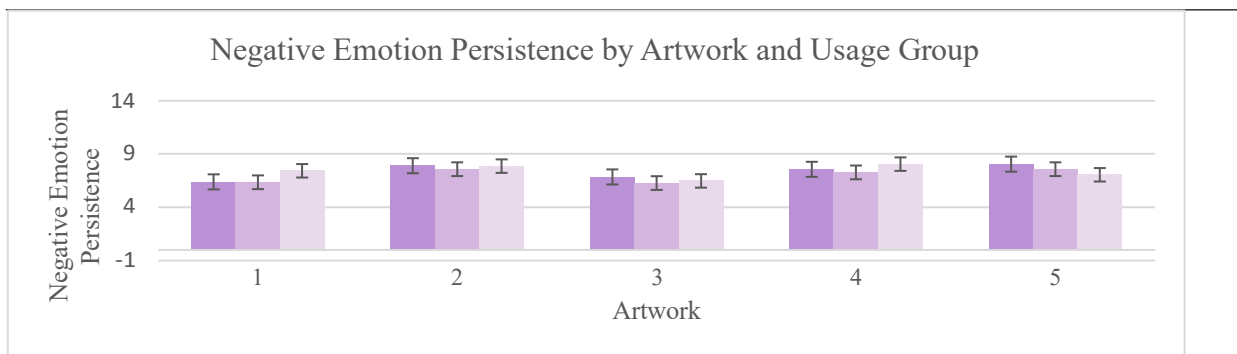


Figure 15. Average persistence of negative emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

Table 10. Average regulation of positive emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
		average	
1	8.939394	8.000000	9.142857
2	7.439394	7.000000	6.761905
3	7.530303	8.150943	7.619048
4	7.015152	7.132075	6.238095
5	8.893939	7.433962	7.619048

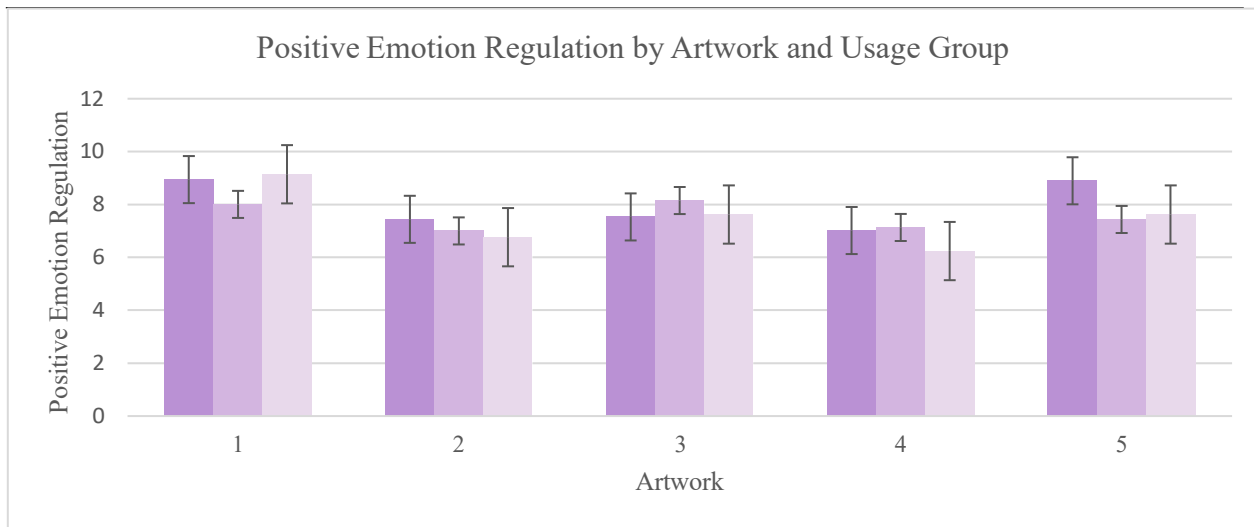


Figure 16. Average regulation of positive emotion response by usage group and artwork. N=140 participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily (n=66); above average = between 127 minutes and 317.5 minutes daily (n=53); significantly above average = above 317.5 minutes daily (n=21). Emotional response assessed by the MEQ.

Table 11. Average regulation of negative emotional response by usage group and artwork.

Artwork #	Below average	Significantly above average	Significantly above average
		average	
1	7.136364	6.54717	7.428571
2	7.863636	7.622642	7.952381
3	7.515152	6.698113	6.571429
4	7.666667	7.54717	7.52381
5	7.772727	7.415094	6.714286

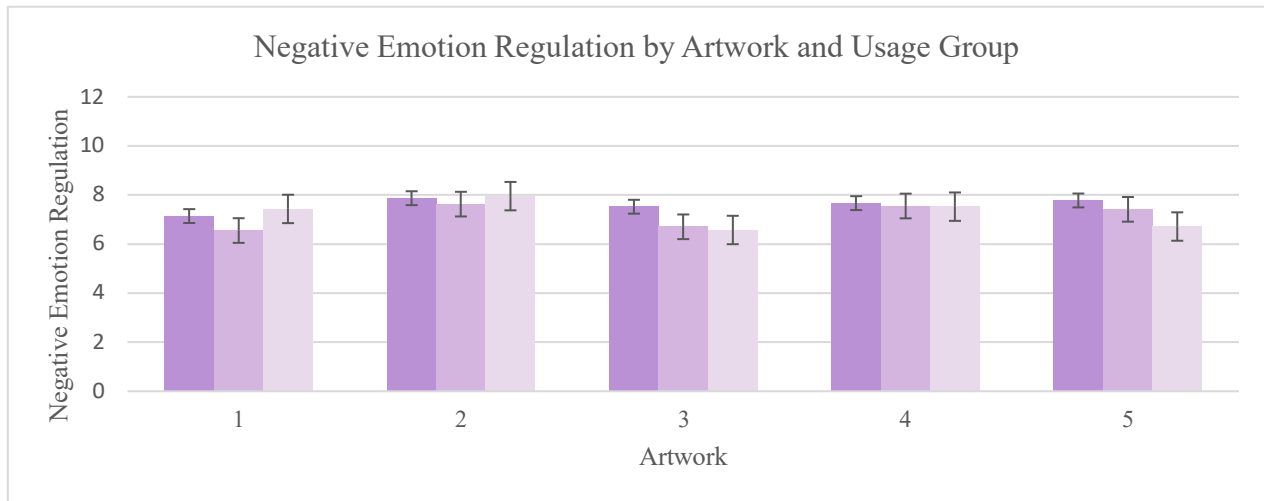


Figure 17. Average regulation of negative emotion response by usage group and artwork. $N=140$ participants split into three different social media usage groups. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

RESULTS

As there were no p-values below 0.05, we fail to reject the null hypothesis. However, trends suggested that although there was no effect of social media usage on negative emotional response, there may be an effect of social media usage on positive emotional response. Overall positive emotional response was greatest in the significantly above average usage group (average= 28.142), followed by the above average usage group (average= 27.170)

and then the significantly above average usage group (average= 26.695). A greater average corresponded to a greater overall positive emotional response. All usage groups had similar average overall negative emotional responses, with the below average usage group having the greatest average (average= 22.736), followed by the significantly above average group (average= 22.667) and then the above average usage group (average=22.128).

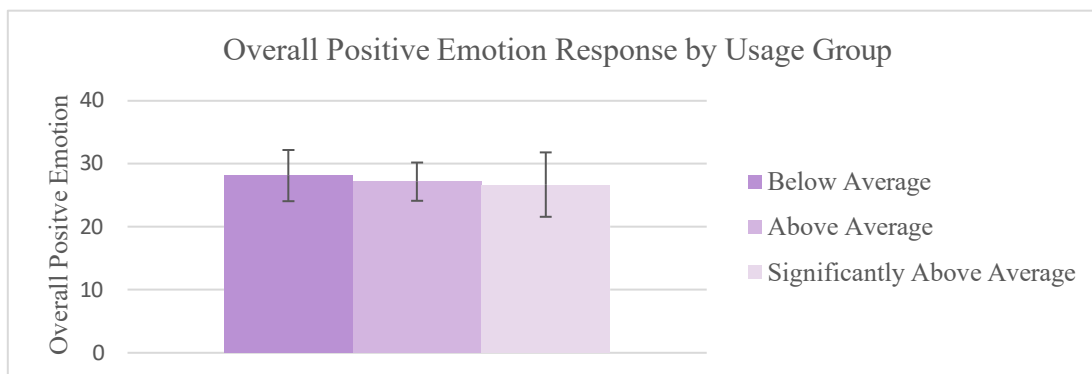


Figure 18. Average overall positive emotional response by usage group across all artworks. Groupings: below average = below 127.5 minutes daily ($n=66$); above average = between 127 minutes and 317.5 minutes daily ($n=53$); significantly above average = above 317.5 minutes daily ($n=21$). Emotional response assessed by the MEQ.

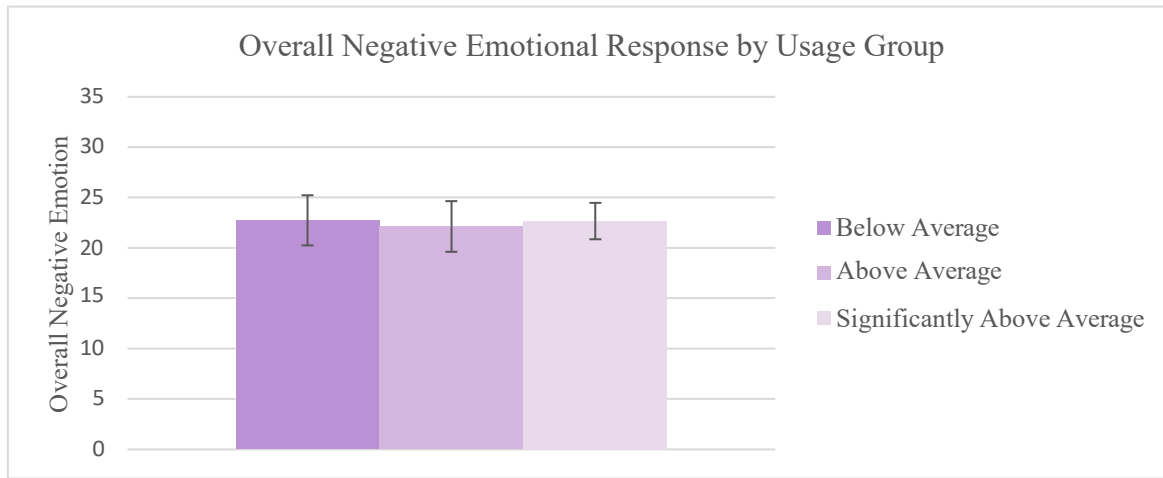


Figure 19. Average overall negative emotional response by usage group across all artworks. Groupings: below average = below 127.5 minutes daily (n=66); above average = between 127 minutes and 317.5 minutes daily (n=53); significantly above average = above 317.5 minutes daily (n=21). Emotional response assessed by the MEQ.

ANOVA testing was done on positive emotional measures, which revealed no statistically significant ($p < 0.05$) differences in emotional responses between social media usage groups.

Table 12. P-values by emotional response from ANOVA testing.

Emotional response (average of all artworks):	Pearson's coefficient (p-value):
Positive overall:	0.860042
Negative overall:	0.90121
Positive frequency:	0.794493
Negative frequency:	0.914081
Positive intensity:	0.886753
Negative intensity:	0.709558
Positive persistence:	0.668083
Negative persistence:	0.626751
Positive regulation:	0.641751
Negative regulation:	0.348313

The most often felt emotional response by artwork was the frequency of happiness for artwork #1 across all three usage groups. Happiness and emotion were the overall most strongly felt emotions, and artworks #1 and #5 had the greatest emotional responses. Further research into prevalence of emotions when observing artworks would need to be conducted to understand these findings in greater detail—for example, are positive emotions always more frequently reported when observing artwork because of potential factors like the social desirability bias, regardless of subject matter? The top three emotional responses for the

below average group were the frequency of happiness for artwork #1 (2.984), intensity of happiness for artwork #1 (2.909), and persistence of happiness for artwork #1 (2.697). The top three emotional responses for the above average group were the frequency of happiness for artwork #1 (2.830), frequency of inspiration for artwork #1 (2.755), and frequency of inspiration for artwork #5 (2.566). The top three emotional responses for the below average group were the frequency of happiness for artwork #1 (2.952), frequency of inspiration for artwork #1 (2.952), and the intensity of inspiration for artwork #1 (2.762).

Table 13. Top five most frequently felt emotional responses for the below average usage (< 127 minutes) group.

Most frequently felt emotional response:	Artwork #:	Value (out of 5):
Happy [How Often?]	1	2.984848
Happy [How Intense?]	1	2.909091
Happy [How Long-Lasting?]	1	2.69697
Happy [How Easy to Regulate?]	1	2.681818
Sad [How Often?]	1	2.681818

Table 14. Top five most frequently felt emotional responses for the above average usage (between 127 minutes and 317.5 minutes) group.

Most frequently felt emotional response:	Artwork #:	Value (out of 5):
Happy [How Often?]	1	2.830189
Inspired [How Often?]	1	2.754717
Inspired [How Often?]	5	2.566038
Happy [How Intense?]	1	2.509434
Happy [How Often?]	3	2.509434

Table 15. Top five most frequently felt emotional responses for the significantly above average usage (> 317.5 minutes) group.

Most frequently felt emotional response:	Artwork #:	Value (out of 5):
Happy [How Often?]	1	2.952380952
Inspired [How Often?]	1	2.952380952
Inspired [How Intense?]	1	2.761904762
Happy [How Long-Lasting?]	1	2.523809524
Happy [How Intense?]	1	2.380952381

Standard deviation was additionally greater for positive emotional responses compared to negative emotional responses across all usage groups. This implies a greater variance in positive emotional response to artworks in general.

CONCLUSIONS

As there were no p-values below 0.05 that would indicate significant differences between social media usage groups' emotional responses to art, we fail to reject the null hypothesis. However, the complexity of the responses gathered were more nuanced than expected when the hypothesis was formed, and there does exist a subtle trend in positive emotional response wherein usage groups with higher social media usage have lower positive emotional response measures.

There is little research investigating social media usage's effect on emotional response to art. The usage of the MEQ to investigate emotion allowed data analysis to investigate not just one value or aspect of emotion, but a detailed method to quantify emotion through five calculated values per emotion—lending itself to a more nuanced resulting analysis and conclusions. However, due to the

novelty of this research, there currently exists no baseline to compare the results of this study to.

Social media and screen time usage was self-reported in this study. Because different mobile platforms will record screen time in different formats (such as cumulative vs. average, by calendar week or past seven days) there may be inaccuracies in user's reports of their own screen time. This risk was minimized by providing direct instructions for finding the correct value (daily average) for different types of phones, but the vast range of user interfaces and cell phone models as well as human error make this issue more complicated.

If it's assumed that individuals will be more willing to purchase and support artwork that they are more emotionally reactive to, current results suggest that it will be more difficult for artists to find potential customers on social media. Although artists may find a relatively more emotionally receptive and willing to buy audience in below average users, promoting on social media will inevitably lead to an audience of mainly above average/significantly above average social media users. However, the potential relationship between emotional reactivity

and support of artwork needs to be further investigated for a formal conclusion to be drawn from this perspective.

APPENDIX

Social Media Usage Statistics

What is your age (in years)? *

Your answer

What is your **average** daily non-work-related screen time over the past **recorded week**? *

*Please include units (hours, minutes)

Your answer

What type of phone do you use? *

☐ Apple/iPhone/iOS
☐ Android/Samsung/Motorola/Google
☐ Other

What is your **daily average** screen time on social media apps over the past **recorded week**? If your device does not calculate the average time for you, **please insert total time used and days recorded** as seen from the mini-calendar. *

(iOS: You can find this recorded/calculated for you by clicking the app icon under Screen Time -> See All Activity -> Week -> Most Used -> app.)

(Android: You can find this recorded/calculated for you in Settings -> Digital Wellbeing and parental controls -> click on the circle diagram, you may need to scroll down.)

Apps that count as social media are: Instagram, Tiktok, Discord, X (Twitter), Facebook, Snapchat, YouTube, Amino, Reddit, etc.

*Please include units (hours, minutes)

Your answer

Back

Next

Page 3 of 11

Clear form

Figure 21. Self-reporting social media and screen time statistics were requested within the survey. The survey was conducted using Google Forms and specific instructions for finding social media and screen time statistics were embedded into the survey.

Artwork 1

Happy *

	1	2	3	4	5
How Often?	None	Once or Twice	Often	Multiple Times Throughout	Nearly Constantly
How Intense?	Very Low	Low	Moderate	High	Very High
How Long-Lasting?	Less than 1 second	1-10 seconds	1 minute	1-3 minutes	Longer than 3 minutes
How Easy to Regulate?	Very Easy	Easy	Moderate	Difficult	Very Difficult

1

2

3

4

5

How Often?

How Intense?

How Long-Lasting?

How Easy to Regulate?

☐

☐

☐

☐

☐

Figure 22. An example of how the artworks appeared alongside the MEQ within the survey. Each of the five artworks were accompanied by the full MEQ.

56

REFERENCES

- Bivins, T. (2021). *Adults are turning to social media apps in bed to avoid emotions*. Vita Health Group. Retrieved December 14, 2023, from <https://www.vitahealthgroup.co.uk/news/young-adults-are-turning-to-social-media-apps-in-bed-to-avoid-emotions/>
- Cheung, M.-C., Law, D., Yip, J., & Wong, C. W. Y. (2019, January 22). *Emotional Responses to Visual Art and Commercial Stimuli: Implications for Creativity and Aesthetics*. *Frontiers*. Retrieved December 13, 2023, from <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00014/full>
- Crossman, A. (2019, June 20). *How to Construct an Index for Research*. ThoughtCo. Retrieved November 12, 2023, from <https://www.thoughtco.com/index-for-research-3026543>
- DHA Communications. (2013, December 3). *Paying Artists research: Phase 1 findings*. a-n. Retrieved January 1, 2024, from <https://www.a-n.co.uk/research/paying-artists-research-phase-1-findings/>
- Doeser, J., TBR's Creative & Cultural Team, & The Artists Information Company. (2018, December 14). *Livelihoods of Visual Artists: Literature and Data Review*. a-n. Retrieved January 1, 2024, from <https://static.a-n.co.uk/wp-content/uploads/2019/01/Livelihoods-of-Visual-Artists-Literature-and-Data-Review.pdf>
- Gausby, A. (2015). *Attention spans*. Consumer Insights, Microsoft Canada. Retrieved January 20, 2024, from <https://dl.motamem.org/microsoft-attention-spans-research-report.pdf>
- Klonsky, D., Victor, S. E., Hibbert, A. S., & Hajcak, G. (2019, June 8). *The Multidimensional Emotion Questionnaire (MEQ): Rationale and Initial Psychometric Properties*. Springer Link. Retrieved December 13, 2023, from <https://www2.psych.ubc.ca/~klonsky/publications/meq.pdf>
- Korte, M. (2020). The impact of the digital revolution ^[L]_[SEP] on human brain and behavior: Where ^[L]_[SEP] do we stand? *Dialogues in Clinical Neuroscience*, 22(2), 101-111. <https://doi.org/10.31887/dcns.2020.22.2/mkorte>
- Lola, O., Juliet, J., & Tan, J. K. (2023, March). *Structurally F-cked an inquiry into artists' pay and conditions in the public sector in response to the Artist Leaks data*. a-n. Retrieved January 1, 2024, from <https://www.a-n.co.uk/research/structurally-f-cked/>
- Moscardo, G., & Hughes, K. (2018, March 12). *Museums Unplugged*. Taylor & Francis Online. Retrieved March 6, 2024, from <https://www.tandfonline.com/doi/abs/10.18666/JLR-2016-V48-I5-7085>
- Pang, C. Y., Nadal, M., Müller-Paul, J.S., Rosenberg, R., & Klein, C. (2012, October 31). *Electrophysiological correlates of looking at paintings and its association with art expertise*. ScienceDirect. Retrieved September 14, 2023, from <https://www.sciencedirect.com/science/article/abs/pii/S0301051112002438>
- Prado, C. G. (2017). *Social media and your brain: How web-based communication is changing how we think and express ourselves*. Praeger.
- Shewale, R. (2024, March 4). *Social Media Users — Global Demographics (2024)*. DemandSage. Retrieved April 22, 2024, from <https://www.demandsage.com/social-media-users/#:~:text=There%20are%204.9%20billion%20social,network%20penetration%20rate%20is%2059.4%25.>
- Statista Research Department. (2023, March). *Number of employees in the arts, entertainment, and recreation industry in the United States from 2013 to 2022, with a forecast for 2023*. Statista. Retrieved April 23, 2024, from <https://www.statista.com/statistics/1180595/number-of-arts-entertainment-recreation-industry-employees-us/>
- Wong, B., J.D. (2023, May 18). *Top Social Media Statistics And Trends Of 2023* (C. Bottorff, Ed.). Forbes Advisor. Retrieved November 1, 2023, from <https://www.forbes.com/advisor/business/social-media-statistics/#:~:text=Social%20Media%20Usage%20Statistics,-The%20average%20person&text=So%20it's%20significant%20that%20the,hours%20and%207%20m inutes%20daily.>

Optimizing 2000m Rowing Performance: An Experimental Study Assessing Ergometer Performance in Relation to Pacing Strategy

Bhanu Atmakuri

ABSTRACT

In rowing, pacing can be the difference between winning and losing a race; can the potentially optimized start of an even pacing strategy in comparison to the current standard parabolic pacing strategy improve the performance of a rower in a 2000m race? The hypothesis was yes, an even pacing strategy will outperform a parabolic pacing strategy by a significant split difference. Six rowers were each instructed to follow one of the two strategies in a 2000m trial: even pacing, which involves maintaining a constant power output until the final 500m, and parabolic pacing (the current standard), which involves a high, lower, and low power output respectively for the first three quarters of the race. For the final 500m of these strategies, participants were instructed to deliver maximum power (similar to the last stretch of an actual race); this was used to determine the validity of the strategies for efficiently distributing power output throughout the race. Although most participants were within two standard deviations of their prescribed power outputs in the first three quarters and the even pacing strategy outperformed parabolic pacing by about 1.3% in the last quarter, the null hypothesis could not be refuted because the difference was not statistically significant ($p = 0.649$). These results, although inconclusive due to a lack of data, suggest that the even pacing strategy could prove to be more viable/efficient in a race setting, and further research must be conducted on the even pacing strategy since the number of trials was inadequate.

INTRODUCTION

This project seeks to compare the even pacing strategy, a strategy designed to stress the body less, against the more commonly adopted parabolic strategy in a 2000m rowing race to ultimately aid coaches better train rowers to race faster. Most often, new rowers are directed to use the same, generally undisputed strategy known as a parabolic strategy: start strong with a fast split time (a dynamic indicator of predicted finishing time based on current power output) then hold a slower consistent split time, and end with as fast a split time as possible. One study confirms that this strategy is universally used throughout Olympic rowers; its

data indicates that Olympic rowers, on average, start at around 700W (W is the rate of Work being done by rowers – another indicator of power) and then decline by the 300m mark to a steady 430-450W range until the 1700m mark, where the rowers end with a gradual incline by about 50W until the end (Das et al., 2019). Although this strategy might seem intuitive because it gets the boat ahead early and conserves the rowers' energy until the end, this strategy has yet to be optimized – which is a big problem for teams aiming to win races. A good example of a potential fault in this strategy is the fact that rowers' heart rates only increase as a race progresses with an average total increase of 8.6% from the start of the race (Garland, 2005).

A potential solution to this problem can be developing a new strategy that addresses the potential faults in this current, universally used strategy.

However, many similar studies that have explored 2000m race strategies have not addressed alternatives to the current standard. One study, similar to this one, has experimented with the same two strategies being tested in this study and found no statistical difference between the two. However, the experiment lacked an adequate number of trials to draw conclusive data – with only two total 2000m trials – and its procedure created confounding variables – such as the verbal encouragement given near the end of the 2000m pieces – which could affect each rower’s performance differently (Clegg, 2020). Apart from this study, others never directly explored the impact of pacing strategies through experimentation. For example, in the same study that examined Olympic rowers’ strategies, the researcher only speculated that the explosive start may cause inhibition of anaerobic glycolysis and muscle contraction, meaning the increased power output may inhibit performance later in the race.

Performance, as defined in this study, is the split time acquired in the last 500m of a 2000m trial. A split time is the time it takes to row 500m and is therefore an indicator of power. More power outputted by a rower means a faster time to row a set distance, meaning that split times and power output are directly correlated. The last 500m of a 2000m race usually includes a sprint, which involves maximum power output. If a strategy has efficiently conserved and distributed power, which the, then the

sprint can be much more powerful. parabolic pacing strategy is speculated to not do

One possible strategy that may avoid the issue with the parabolic pacing strategy is the even pacing strategy, where the rower holds a steady split time until the end and similar to a parabolic strategy, ends with as fast a split time as possible. This strategy aims to eliminate the problem-causing explosive start and can potentially create a much more gradual increase in metabolic acidosis and heart rate instead of a sharp increase. As a result, overall inhibition of anaerobic glycolysis/muscle contraction is lessened in comparison to the parabolic pacing strategy. As a result, this “optimized” start could result in the even pacing strategy outperforming the parabolic pacing strategy, which was the hypothesis for this study.

METHODOLOGY

Participants were sampled by visiting the local rowing clubs that has around 40-50 male experienced (with 1 year of experience as a high school rower) rowers. In total, 10 participants were sampled. However, 4 of them decided not to participate later or during the study due to not wanting to perform the exerting task of a 2000m trial, sicknesses, or other personal reasons. This study only sampled for male high school rowers because middle schoolers or collegiate rowers would not generate the same amount of power or consistency (middle schoolers would produce less and collegiate would produce more) than high school rowers. Similarly, females would also differ in power output. The coaches/club heads were asked for permission to announce a request for

participation in the study and to post flyers around the club area (boathouse, nearby poles, etc. Consent forms were provided to people wanting to participate. Names were collected in order to organize data with the random ID values and were later discarded once all data was collected and correctly paired with the participants ID values in a spreadsheet document. Other metrics that were collected include PR (personal record/best) split times and other physical attributes: age (yrs), years of experience (yrs), and PR split time for a 2k ergometer (rowing machine) piece (min). These metrics were stored in Data Table #1. In this experiment, a Microsoft sign-up form was used to collect names, PR split times, and physical attributes for each participant.

After sampling, a random order of strategies was assigned for each participant for all the pieces they did. This is to ensure that increasing performance of the rowers due to training over the testing period does not result in one strategy outperforming the other. Before the pieces, rowers were instructed 48 hours in advance (through messaging) to eat healthy carbohydrates like bread, noodles, fruits, and green vegetables (quickly digesting and energizing foods – high in carbs, low in fibers and fat preferably) to fuel for the trial and

to drink lots of water to stay hydrated. Additionally, they were instructed to get the recommended amount of at least 8 hours of sleep for each night. This may help reduce the confounding effects of varying energy levels and mental capacity during trials. The rowers had access to three sheets: one with their assignment of strategy for each piece, another with any other instructions about how to set up an ergometer with RowHero, and one with what to do before the piece (nutrition/other guidelines as mentioned in the previous sentences). These were provided through text and in person before the trials.

Flowchart: Random Assignment of Strategy

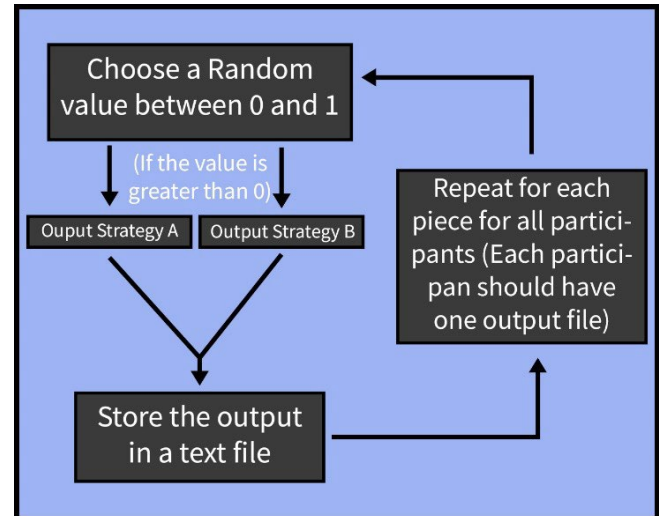


Figure 1. Randomizes chooses strategy A or B for each participant and creates an output file with ID's and strategies with that random decisions. Repeats for all pieces.

Example Split Goal Instruction sheet:

Participant ID:	Strategy	Goal Split: first 500m	Goal Split: second 500m	Goal Split: third 500m	Goal Split: fourth 500m	Goal Average Split:
1	A	Ex: 1:45.8	Ex: 1:46.2	Ex: 1:46.8	Ex: 1:46.3	Ex: 1:46.2
2	B	Ex: 2:03.4	Ex: 2:03.9	Ex: 2:04.6	Ex: 2:04.0	Ex: 2:04.1
Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.

Figure 2A. This is an example of the goal splits participants were shown and instructed to follow.

Example Additional Instructions sheet:

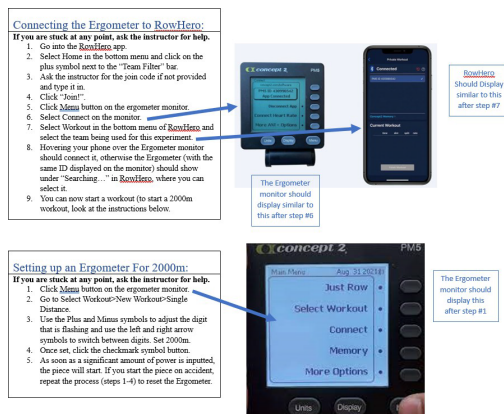


Figure 2B. The top left box describes how to connect RowHero (the application used to record data into from the ergometers) to the ergometers. The bottom left box describes how to set up a 2000m trial on an ergometer (RowHero 2023).

Example Nutrition Guidelines sheet:

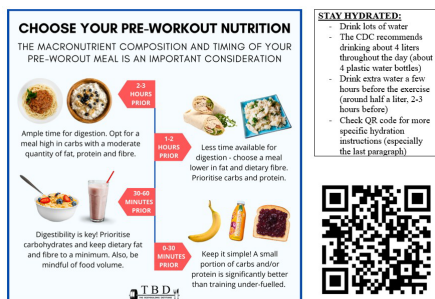


Figure 2C. This sheet describes basic nutritional advice in preparation for a physical workout (Nelson, 2023 and BHF) and gives information on hydration (Kaehler, 2011)

Goal split times were calculated using the equation $100SSSSSSS$ (Sp – split time, P – percent velocity).

This equation gives the split relative to the split time inputted when given a percent velocity.

Before starting the pieces, participants connected their ergometers to the application (RowHero) being used to record the data and set up for a 2000m piece. The data was exported to excel at

the end through the app and recorded into Data Table #3. Each participants performed two total 2000m pieces which were spaced by at least four days over nine weeks. One participant was unable to perform their second trial due to sickness. An algorithm that takes a list of participants and assigns them an identification number was used to replace and discard names after data collection.

Data Table #1 – Participant Physical Metrics

Participant ID Number	Age (yrs)	Experience (yrs)	PR (min)
1			
2			
3			

Figure 4. This data table holds each participant's metrics. This data has been discarded.

Flowchart: Random Assignment of ID

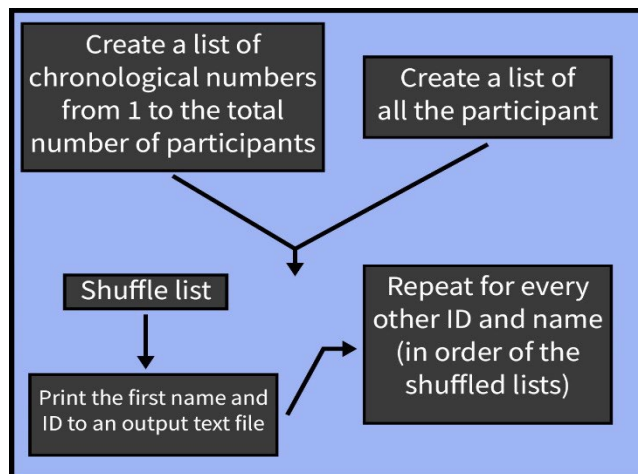


Figure 3. Randomizes order of ID's (1 to number of participants) and creates an output file with names and ID's with that random order.

Data Table #2 –Overall Performance

Participant ID	Q1	Q2	Q3	Q4
1 (B)	99.19	100.54	104.4	103.01
2 (B)	101.7	101.09	100.3	95.598
3 (A)	96.14	101.21	100.9	103.41
4 (B)	97.59	101.22	102	99.172
5 (A)	98.87	98.963	100.9	99.564
6 (A)	99.33	100.49	101.9	98.632
1 (B)	100.7	101.46	103.9	105.51
2 (A)				
3 (B)	99.29	99.335	100.7	98.911
4 (A)	97.63	99.338	98.89	100.68
5 (B)	98.49	99.532	100.3	97.609
6 (A)	99.4	100.06	101.6	99.253

Figure 4B. This data table holds the strategy performance data for each trial along with the participant ID of the person who performed the trial. Performance data is the split time divided by goal split time (this produces a percent) from the last 500m of the 2000m trials.

Data Table #3 – Ergometer Data (for each participant)

Distance (m)	Split time (sec)

Figure 4C. This data table represents the raw excel data for each

Data Table #4 – ST DEV Analysis (Even Pacing)

STDEVAnalysis	Q1	Q2	Q3	Mean	2STD	1STD
(A)						
3 (A)	96.13636364	101.2121212	100.9032634	99.41724942	5.691048029	2.845524
5 (A)	98.86561955	98.96257514	100.8629048	99.56369983	2.252376746	1.126188
6 (A)	99.32532348	100.4898336	101.922366	100.5791744	2.601648538	1.300824
4 (A)	97.63205829	99.33763868	98.89146474	98.62038724	1.769026146	0.884513
6 (A)	99.39835916	100.0638104	101.6499544	100.370708	2.313490763	1.156745

Figure 4D. This data table holds the standard deviation analysis for strategy A (Even Pacing). All quarters fell within 2 standard deviations, and most (ones not highlighted in yellow) fell within 1 standard deviation.

Data Table #5 – ST DEV Analysis (Parabolic Pacing)

STDEVAnalysis (B)	Q1	Q2	Q3	Mean	2STD	1STD
1 (B)	99.18795801	100.5351236	104.3923159	101.3717991	5.402353071	2.701177
2 (B)	101.6553067	101.0855683	100.2532447	100.9980399	1.410234583	0.705117
4 (B)	97.58677346	101.2222222	101.9637311	100.2575756	4.685017333	2.342509
5 (B)	98.49424269	99.53220191	100.3346287	99.4536911	1.845403075	0.922702
1 (B)	100.7130125	101.4594279	103.9386041	102.0370148	3.377167625	1.688584
3(B)	99.29015084	99.33522232	100.7028547	99.77607594	1.605860474	0.80293

Figure 4E. This data table holds the standard deviation analysis for strategy B (Parabolic Pacing). All quarters fell within 2 standard deviations, and most (ones not highlighted in yellow) fell within 1 standard deviation.

RESULTS/DISCUSSION

The hypothesis, that the even pacing strategy will outperform the parabolic pacing strategy, has been refuted. With outliers included, there was almost no difference in performance between the strategies with a difference of 0.33% ($p = 0.8491$). Without the outliers, however, the difference was about 1.8% ($p = 0.2181$), and the p value suggests that there might be merit to this difference, but by conventional standards, the difference is not statistically significant. The results of this study therefore suggest that both strategies perform equally well, however, this result may not be representative of the larger population because the data is not very precise.

For example, many abnormal patterns had occurred in the data such as some trials where participants started much faster than they were instructed to and then went slower in the following quarters likely as a result of that excessive power output at the beginning. Because there has not been scientific precedent on this matter, a threshold percentage was chosen based on the observed minimum deviation of trials of which these

abnormal patterns. To remove these outliers, threshold percentages were used where $\pm 2\%$ from the given goal split was used to check that all the participants performed each quarter close enough to their instructed splits that the performance of the trial reflected the strategy. 7 trials (3 from the even pacing strategy, which had a total of 6 trials, and 4 from the parabolic pacing strategy, which had a total of 7 trials) were deemed outliers due to not meeting the threshold. Standard deviation analysis revealed that each of the first 3 quarters fell within 2 standard deviations away from the mean. Although, a few of the first 3 quarters from trials of both strategies did not fall within 1 standard deviation (Figure 4D and 4E). Both strategies had trials with abnormally high standard deviations, and most of these trials coincided with the outliers found earlier.

The biggest limitation in this study is a lack of data and sufficient trials. Outliers with the high standard deviations can likely be attributed to this lack of sufficient trials because participants were still unfamiliar with the strategies they performed as of the second trial. Although both strategies were performed within statistical bounds, they were not

performed as precisely as intended. The lack of data may also contribute to means that are less representative of the larger population. This means

that more trials with the same participants (to build familiarity with the strategies) and more data would be the next steps for this study.

BIBLIOGRAPHY

- Boillet A, Haas B, Samozino P, Morel B, Bowen M, Cohen C, Messonnier LA. Is the Most Commonly Used Strategy for the First 1,500 m of a 2,000 m Rowing Ergometer Race the Most Appropriate? *Front Physiol.* 2022 Mar 8;13:827875. doi: 10.3389/fphys.2022.827875. PMID: 35350683; PMCID: PMC8958042.
- British Heart Foundation. (n.d.). *Food for fitness: What should I eat before a workout?*. BHF. <https://www.bhf.org.uk/information-support/heart-matters-magazine/nutrition/food-for-fitness#:~:text=Choose%20something%20that%20includes%20carbohydrates,cause%20stomach%20discomfort%20during%20exercise>
- Chu, D., Tsai, M.-C., Sheehan, R., Davis, J., & Doig, R. (2023, January 1). *Identifying pacing profiles in 2000 metre World Championship Rowing*. *Journal of Sports Analytics*. <https://content.iospress.com/articles/journal-of-sports-analytics/jsa220497>
- Clegg, R. A. (2020, September 11). *The effect of pacing strategy on 2000 m rowing ergometer performance in well-trained male and female rowers*. Western Libraries. <https://ir.lib.uwo.ca/etd/7483/>
- Connecting to the PM5*. RowHero. (2023, February 25). <https://rowhero.com/help/getting-started-for-rowers/connecting-to-the-pm5/>
- DAS, A., Mandal, M., Syamal, A. K., & Majumdar, P. (2019). Monitoring Changes of Cardio-Respiratory Parameters During 2000m Rowing Performance. *International journal of exercise science*, 12(2), 483–490.
- Garland S. W. (2005). An analysis of the pacing strategy adopted by elite competitors in 2000 m rowing. *British journal of sports medicine*, 39(1), 39–42. <https://doi.org/10.1136/bjsm.2003.010801>
- Kaehler, B. (2011, April 5). *Drink up! hydration and training*. row2k.com. <https://www.row2k.com/features/541/drink-up--hydration-and-training/>
- Martins, A. N., Borges, T. O., Massa, M., Bacurau, R. F., Moreira, A., & Aoki, M. S. (2022). Pacing Strategy of Single Scullers during Rowing World Championships. In *IJASS(International Journal of Applied Sports Sciences)* (Vol. 34, Issue 2, pp. 215–224). Korea Institute of Sport Science. <https://doi.org/10.24985/ijass.2022.34.2.215>
- Nelson, T. (2023, March 12). *What you should eat before a workout*. The Bodybuilding Dietitians. <https://www.thebodybuildingdietitians.com/blog/what-you-should-eat-before-a-workout>

Testing the effectiveness of progressive muscle relaxation on the cognitive performances of adolescents with somatic anxiety.

Inaya Chowdhury

ABSTRACT

Progressive muscle relaxation is a common tool used for anxiety reduction, but is it effective at improving the performance of the anxious brain? Somatic, or physically impairing, symptoms are highly prevalent among children and adolescents with anxiety disorders and are associated with reported structural change and functional disturbance of the brain. This study seeks to address the rapidly growing incidence of somatic anxiety in our youth and its alarming consequence—cognitive decline—through the implementation of progressive muscle relaxation training.

Fifteen participants were placed in either a control, experimental, or alternate experimental group depending on their quantified levels of somatic anxiety measured by the Cognitive-Somatic Anxiety Questionnaire. All participants engaged in varying amounts of guided progressive muscle relaxation (independent variable) and completed a series of cognitive tests to receive an accumulated score of cognitive performance (dependent variable). The experimental hypothesis was that greater durations of progressive muscle relaxation would improve the cognitive performance of adolescents with somatic anxiety.

The results of the investigation supported the hypothesis with the greatest mean difference in test scores from Trial 1 to Trial 2 being of the experimental group (6.7 points), far greater than the mean differences of the other 2 groups. A one-way ANOVA—analysis of variance—confirmed a statistically significant relationship between minutes of progressive muscle relaxation and improved test scores among all groups, with a p-value of approximately 0.04. This study introduces a possibly beneficial intervention for somatic symptoms accompanied by debilitating conditions beyond anxiety disorders.

INTRODUCTION

This study seeks to address the rapidly progressing incidence of somatic anxiety in our youth and its alarming consequence—cognitive decline—by the implementation of progressive muscle relaxation training. Somatic, or physical, symptoms are highly prevalent among children and

adolescents with anxiety disorders and are associated with greater anxiety severity and impairment, including restlessness (74%), stomachaches (70%), and palpitations (48%) as some of the more common symptoms (Ginsburg, Riddle, & Davies, 2006). A more recent investigation has furthered research on a particular effect of anxiety impairment: some

Figure 1. Mean Perceived Stress (PSS), State Anxiety (STAI), Working Memory Capacity (TMTB), and Academic Performance (PRACT) Scores

Assessment and Group	Pretest Score, <i>M</i>	Posttest Score, <i>M</i>
PSS		
Control group	17.41	18.36
Experimental group	17.22	16.44
Total	17.31	17.40
STAI		
Control group	40.66	40.88
Experimental group	39.80	33.17
Total	40.23	37.02
TMTB		
Control group	44.83	41.34
Experimental group	47.45	37.91
Total	46.14	39.63
PRACT		
Control group	85.94	83.59
Experimental group	83.95	90.64
Total	84.95	87.12

Figure 1. displays the positive effect of PMR on the state anxieties, working memories, and academic performances of participants in the experimental group.

Source: Hubbard, K. K., & Blyler, D. (2016). *Improving Academic Performance and Working Memory in Health Science Graduate Students Using Progressive Muscle Relaxation Training*. *American Journal of Occupational Therapy*, 70(6), 7006230010p1.

reported structural change and functional disturbance of the brain in patients with bodily symptoms (Fu et al., 2019). These data suggest there may be a higher likelihood of cognitive impairment in people carrying a greater number of somatic symptoms.

Edmund Jacobson's progressive muscle relaxation (PMR), developed in the early 1920's, is a commonly used relaxation technique that reduces stress and anxiety in the body through the process of tensing and releasing muscle groups and deep breathing. This technique can be used to treat both cognitive anxiety, when anxiety symptoms tend to occur more so within the brain, and somatic anxiety, when anxiety symptoms tend to occur more so physically in the body.

Numerous studies have been conducted on the extended benefits of progressive muscle relaxation beyond relieving anxiety—useful precursors to exploring the extended benefit of PMR in improving cognitive performance. One experiment, conducted in 2016, tested the effect of PMR training on graduate students to determine whether it could result in a measurable decrease in a stressful environment surrounding their practical examinations (Hubbard, Blyler, 2016). The experiment used a convenience sample of 128 health science graduate students at the University of St. Augustine. Participants were randomly assigned to the experimental PMR group or the control group.

Data was collected through scaling instruments—the Perceived Stress Scale, the State-Trait Anxiety Inventory, and the Trail Making Test (Part B)—and from administered practical examinations. Pretest data collection occurred in the fourth week of the semester before any practical examinations. Subsequently, in weeks 5-8, four-session PMR training took place and lasted approximately 20 minutes each session. Data was analyzed with a repeated-measures analysis of variance (ANOVA) comparing the mean score changes from pretest to posttest differences between the two groups. The results of the experiment are displayed in *Figure 1*. The findings indicate a positive effect of PMR on the state anxiety of students in the experimental group, with a reduction in mean inventory scores of approximately 6 from pretest to posttest, $p < .001$. In addition, PMR training had a major impact on the academic performance of the students, as there was a large difference in the means between the pretest (84.96)

and the posttest (87.12), $p < .001$. Overall, PMR training had a significant positive effect on the state anxiety, working memory capacity, and academic performance of the students—hinting at the technique’s large range of beneficial effects.

Another experiment conducted in 2016 studied the effect of a highly truncated relaxation intervention, in the form of an 11-minute progressive muscle relaxation exercise, on derived relational responding, a field regarding stimulus-related tasks (Tyndall, Howe, Roche, 2016). In the investigation, 35 participants were exposed to matching-to-sample training to establish a baseline. They were subsequently placed in a PMR group or one of two control groups: simple or conditional discrimination task. During the five phases of the procedure, participants were randomly assigned to one of three experimental conditions while sitting in front of a computer screen displaying instructions on each task. In phase 4 of interventions, the PMR group listened to recorded PMR instruction sets based on Jacobson’s (1938) principles, the simple discrimination task group was instructed to click on only red circles in a series of blue and red circles for their task, and the conditional discrimination task group was instructed to repeat a matching-to-sample task from phase 1—baseline acquisition. Of the 35

participants, those in PMR intervention scored significantly higher mean response accuracies for total ($M=27.83$), one-node ($M=13.50$), and two-node ($M=15.17$) probe responses. One-node probe responses refer to simpler derived relational responding tasks than two-node probe responses. Overall, results were indicative of positive effects of PMR intervention on derived relational responding tasks—an area of focus that falls under cognitive function, which is useful background for a novel study on overall cognitive performance.

A separate experiment, carried out in 2015, investigated the effects of two PMR training frequencies—6 and 12 sessions—on emotional distress, sustained attention, and short-term memory in 10- and 11-year-old school children (Hashim, Zainol, 2015). The sample size of participants was 132 primary school students. The experiment used a 30-minute prerecorded PMR instructive CD, the Depression Anxiety Stress Scale-21 (DASS-21), the digit span test (DST), and the Digit Vigilance Test (DVT). Participants were randomly assigned to one of the experimental groups, 6 or 12 sessions of PMR training, or the control group. The relaxation groups engaged in the PMR intervention during Health Education class, while the control group engaged in its normal Health Education lesson.

Figure 2. Mean $\pm$ Standard Deviations for the Primary Measures at Pre- and Post-intervention by Group

	6-Session group	12-Session group	Control group
<i>Pre-intervention</i>			
Stress	11.90 $\pm$ 7.65	11.57 $\pm$ 7.55	9.35 $\pm$ 6.94
Anxiety	8.24 $\pm$ 6.91	8.42 $\pm$ 7.16	6.70 $\pm$ 4.59
Depression	9.26 $\pm$ 7.3	7.96 $\pm$ 6.13	4.86 $\pm$ 5.13
Short-term memory	9.87 $\pm$ 1.96	8.40 $\pm$ 2.14	10.39 $\pm$ 1.7
Sustained attention	28.778 $\pm$ 4.13	34.30 $\pm$ 8.1	31.45 $\pm$ 5.810
<i>Post-intervention</i>			
Stress	10.58 $\pm$ 9.49	11.19 $\pm$ 10.42	8.54 $\pm$ 5.8
Anxiety	9.75 $\pm$ 8.26	8.17 $\pm$ 8.3	7.29 $\pm$ 5.7
Depression	9.65 $\pm$ 9.83	8.56 $\pm$ 9.2	5.40 $\pm$ 6.3
Short-term Memory	10.84 $\pm$ 2.24	10.26 $\pm$ 2.49	10.86 $\pm$ 2.29
Sustained Attention	35.21 $\pm$ 5.95	38.45 $\pm$ 9.07	36.90 $\pm$ 3.64

Figure 2. displays the substantial positive effect of PMR on the short-term memories of participants in the 12-session group.

Source: Hashim, H. A., & Zainol, N. A. (2015). Changes in emotional distress, short term memory, and sustained attention following 6 and 12 sessions of progressive muscle relaxation training in 10–11 years old primary school children. *Psychology, Health & Medicine*, 20(5), 623–628.

Figure 3. CSAQ Scale

Item No. and content ^a
1. I find it difficult to concentrate because of uncontrollable thoughts. ^c
2. My heart beats faster.
3. I worry too much over something that doesn't really matter.
4. I feel jittery in my body.
5. I imagine terrifying scenes.
6. I get diarrhea.
7. I can't keep anxiety-provoking pictures out of my mind.
8. I feel tense in my stomach.
9. Some unimportant thought runs through my mind and bothers me.
10. I nervously pace.
11. I feel like I am losing out on things because I can't make up my mind soon enough
12. I become immobilized.
13. I can't keep anxiety-provoking thoughts out of my mind.
14. I perspire.

Figure 3. displays the Cognitive-Somatic Anxiety Questionnaire. Round 1 used this scale to place participants in Groups A, B, or C. Responses range from 1 ("not at all") to 5 ("very much so").

Source: DeGood, D. E., & Tait, R. C. (1987). The cognitive-somatic anxiety questionnaire: Psychometric and validity data. *Journal of Psychopathology and Behavioral Assessment*, 9(1), 75–87.

The results of an ANOVA (variance) revealed a significant difference in mean changes in short-term memory ($F=3.89$, $df=2$, $p < .05$)

between the two groups. The highest memory increase was observed in the 12-session group. However, no significant differences in anxiety ($F=.56$, $df=2$, $p > .05$), depression ($F=.52$, $df=2$, $p > .05$), stress ($F=.11$, $df=2$, $p > .05$), or sustained attention ($F=.11$, $df=2$, $p > .05$) occurred between the groups.

With these three studies serving as significant background research and additional preceding investigations on the association between developmental impairment and somatized conditions, the novel study outlined below sought to build upon existing knowledge by establishing a causation relationship between progressive muscle relaxation and different areas of cognitive performance—such as attention, visuospatial skills, and verbal memory—of adolescents who demonstrated a high indication of somatic symptoms of anxiety. Subsequently, the effectiveness of progressive muscle relaxation was tested on the cognitive performances of subjects who demonstrated a high indication of cognitive symptoms of anxiety for comparison.

The hypothesized results of this study were indicative of greater improvements in the cognitive performances of adolescents with somatic anxiety following increased durations of progressive muscle relaxation.

METHODOLOGY

Fifteen adolescents (aged 10-19) were obtained from Washington state via sending out flyers. Each participant was assigned a random number; during the investigation, participants

were only referred to by their designated numbers. During Round 1, a Microsoft Forms survey containing the Cognitive-Somatic Anxiety Questionnaire (CSAQ) was sent out to each participant over email. Once all 15 responses were received, participants with a higher somatic reading than cognitive reading were put into one of two groups using a random number generator:

- Group A: the control group undergoing no progressive muscle relaxation.
- Group B: the experimental group undergoing varying durations of progressive muscle relaxation.

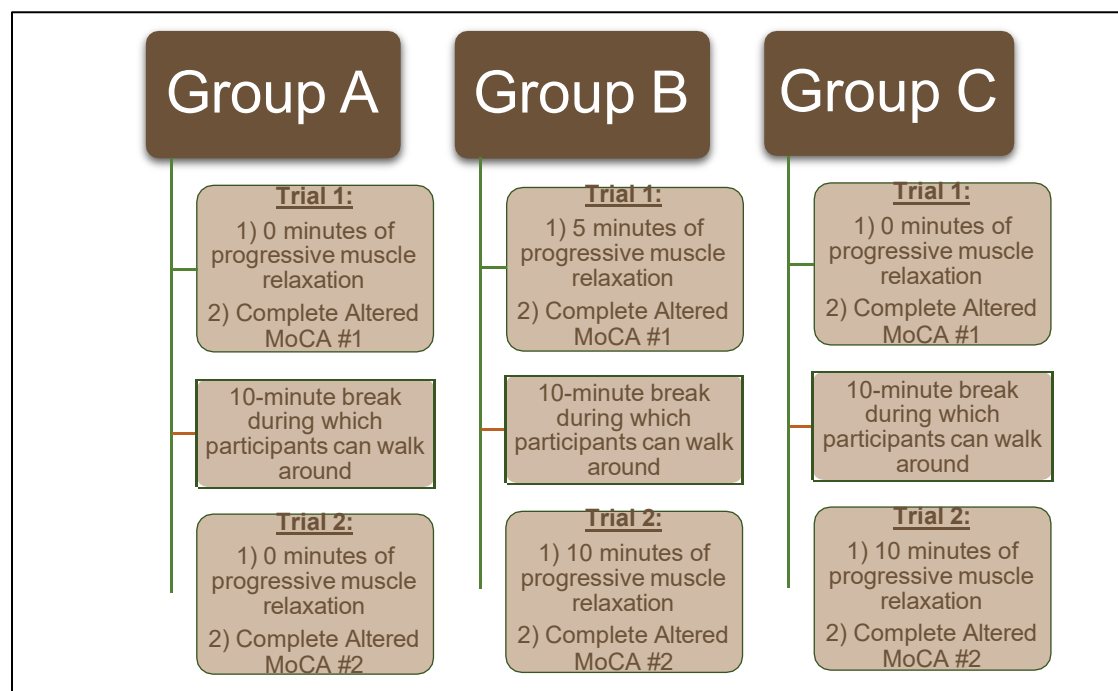
Participants with a higher cognitive reading than somatic reading were put into Group C, an

alternate experimental group that performed varying durations of progressive muscle relaxation to compare with the main experimental group.

Round 2 took place in one-on-one meetings in library meeting rooms or in online meetings. During two trials of experimentation, participants performed two different series of 5 cognitive tests, entailing different aspects of the Montreal Cognitive Assessment (MoCA): language fluency, working and short-term memory, attention and concentration, delayed recall, and visuospatial skills. The original MoCA could not be used as it is tailored towards an older population.

Figure 4. Experimental Procedure Setup

Figure 4. depicts the experimental procedure. Round 2 proceeded with two altered versions of the Montreal Cognitive Assessment (MoCA) and progressive muscle relaxation.



Round 2 Tests

The language fluency test was taken from the original MoCA because its difficulty was viable for an adolescent population. In this test, participants were given one minute to name as many words as they could within a minute. The number of words spoken by each participant was recorded.

The Digit Span test was taken from a free, online website. In this test, 9 series of numbers were verbalized and displayed on the screen (each subsequent series increasing in the number of digits by one) and participants had to type each series in response to the best of their memories. There were additionally 8 separate series which called for participants to type the reverse sequences of each series. The number of series each participant could recall was recorded.

The attention and concentration test was taken from a free, online website. In this test, 8 sequences of differently arranged squares quickly flashed on the screen, before panning to a second sequence. Each participant had to decide whether

the second sequence was the same or different as the first. The number of sequences each participant answered correctly was recorded.

The delayed recall test was taken from the original MoCA for the same reason as the language fluency test. In this test, participants were read 5 words at the beginning of the experiment, which they had to recall 15 minutes later with no cues (cues were given after a certain duration of no response). The number of words each participant was able to recall, as well as how many cues they used, were recorded.

The visual reasoning test was taken from a free, online website. In this test, 15 figures were displayed on the screen, which the participants had to match with their unfolded or folded counterparts. The number of figures each participant matched correctly was recorded.

Once all data was gathered in the data tables, data analysis was performed with mean, median, and mode calculations, as well as a one-way ANOVA in Microsoft Excel.

*Data Table 1.1. Somatic and Cognitive Readings*Scores are between 7-35 points for both scales.*

<i>Participant Number</i>	<i><u>Somatic Scale Reading</u></i>	<i><u>Cognitive Scale Reading</u></i>
<i>1</i>	17	14
<i>2</i>	17	29
<i>3</i>	25	24
<i>4</i>	12	14
<i>5</i>	13	18
<i>6</i>	11	10
<i>7</i>	12	21
<i>8</i>	19	22
<i>9</i>	24	20
<i>10</i>	22	17
<i>11</i>	26	25
<i>12</i>	20	17
<i>13</i>	22	20
<i>14</i>	18	17
<i>15</i>	24	22

Data Table 1.2. Trial 2 Data Table

**The language fluency and delayed recall tests are taken from the original MoCA assessment.*

Participant Number	Digit Span	Spatial Reasoning	Language Fluency	Attention & Concentration	Delayed Recall	<u>Sum</u>
Group A						
1	5 / 17	3 / 15	15	8 / 8	15 / 15	46 / 56
2	8 / 17	2 / 15	20	7 / 8	10 / 15	47 / 56
3	4 / 17	2 / 15	15	7 / 8	14 / 15	42 / 56
4	5 / 17	2 / 15	19	8 / 8	12 / 15	46 / 56
Group B						
7	5 / 17	2 / 15	16	7 / 8	13 / 15	43 / 56
8	7 / 17	3 / 15	19	7 / 8	14 / 15	50 / 56
9	4 / 17	3 / 15	19	8 / 8	15 / 15	49 / 56
10	7 / 17	0 / 15	21	8 / 8	15 / 15	51 / 56
11	4 / 17	2 / 15	18	7 / 8	15 / 15	45 / 56
12	7 / 17	3 / 15	21	7 / 8	12 / 15	50 / 56
Group C						
13	6 / 17	2 / 15	20	7 / 8	12 / 15	47 / 56
14	7 / 17	2 / 15	15	6 / 8	15 / 15	45 / 56
15	6 / 17	0 / 15	7	8 / 8	13 / 15	34 / 56
16	7 / 17	4 / 15	17	7 / 8	15 / 15	50 / 56
17	3 / 17	1 / 15	14	8 / 8	14 / 15	40 / 56

RESULTS AND CONCLUSIONS

Near-equal proportions of participants in each group were achieved. Among all groups, Group B (experimental) had the greatest mean increase in cognitive test scores from Trial 1 to

validity of the spatial reasoning test, due to the highest score being a 3/15; and confounding variables, which included demographics of the participants and lifestyle choices. Lifestyle choices that may have impacted the results of this

Trial 1							
Group	Human #	MoCA	Lang.	Visuo.	Digit	Attn.	Recall
B	1	33	10	1	3	8	11
C	13	49	20	2	4	8	15
B	28	45	15	3	7	7	13
C	16	42	13	2	6	7	14
C	31	38	8	3	7	5	15
A	24	49	24	1	2	7	15
C	4	53	23	2	6	8	14
C	32	35	12	1	3	7	12
B	8	38	15	2	3	7	11
A	26	43	18	2	4	7	12
B	19	47	19	3	4	8	13
A	12	38	12	1	5	7	13
B	3	41	16	2	4	6	13
A	22	40	14	3	3	8	12
B	30	44	17	2	6	8	11

Trial 2					
MoCA	# Words	Visuo.	Digit	Attn.	Recall
43	16	2	5	7	13
47	20	2	6	7	12
50	19	3	7	7	14
45	15	2	7	6	15
34	7	0	6	8	13
46	15	3	5	8	15
50	17	4	7	7	15
40	14	1	3	8	14
49	19	3	4	8	15
47	20	2	8	7	10
51	21	0	7	8	15
42	15	2	4	7	14
45	18	2	4	7	15
46	19	2	5	8	12
50	21	3	7	7	12

Figure 6. Trial 1 and 2 Results

Figure 4. depicts the raw data collected from Trials 1 and 2. Each participant's scores on every cognitive test were added to generate a total altered MoCA score for the participant.

Trial 2 with an average of 6.7. Group B also had the greatest median increase in cognitive test scores with an increase of 7. A one-way ANOVA test, shown below, was conducted in Microsoft Excel on a table of trial differences for each group and calculated a statistically significant p-value of approximately 0.04. The null hypothesis was rejected, and the original hypothesis was therefore accepted: increased durations of progressive muscle relaxation improved the cognitive performances of adolescents with somatic anxiety.

Limitations of this study were the sample size, as 15 participants with little variation in gender, race, and age were used; the predictive

study are the health of the participants and their experiences with relaxation techniques. The low number of trials may have also played a role in skewing the results of this study.

Future steps include conducting further research to validate the relationship between somatization and cognitive impairment, increasing the number of trials in a rerun experiment, ensuring a greater representative sample, and utilizing different cognitive tests with a greater number of cognitive domains. A future study would benefit from using brain scan technology to observe brain activity during and after PMR sessions.

BIBLIOGRAPHY

- Fu, X., Zhang, F., Liu, F., Yan, C., & Guo, W. (2019, February 27). Editorial: Brain and somatization symptoms in psychiatric disorders. *Frontiers*.
<https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00146/full>
- Hashim, H., & Zainol, N. (2015). Changes in emotional distress, short term memory, and sustained attention following 6 and 12 sessions of progressive muscle relaxation training in 10-11 years old primary school children. *Psychology, Health & Medicine*, 20:5, 623-628.
- Hubbard, K., & Blyler, D. (2016). Improving Academic Performance and Working Memory in Health Science Graduate Students Using Progressive Muscle Relaxation Training. *American Journal of Occupational Therapy*, 70, 7006230010.
- M., G. G. M. (n.d.). Somatic symptoms in children and adolescents with anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*.
<https://pubmed.ncbi.nlm.nih.gov/17003663/>
- Montreal Cognitive Assessment (MOCA). (n.d.-a). <https://www.parkinsons.va.gov/resources/MOCA-Test-English.pdf>
- The cognitive-somatic anxiety questionnaire: Psychometric and Validity Data. (n.d.).
https://www.researchgate.net/profile/Raymond-Tait/publication/226652970_The_cognitive-somatic_anxiety_questionnaire_Psychometric_and_validity_data/links/5807deb408aefaf02a2c645c/The-cognitive-somatic-anxiety-questionnaire-Psychometric-and-validity-data.pdf
- Tyndall, I., Howe, B., & Roche, B. (2016). Exposure to Progressive Muscle Relaxation Leads to Enhanced Performance on Derived Relational Responding Tasks. *Association for Behavior Analysis International 2016*.

BIOMEDICAL, CELLULAR BIOLOGY

Autotoxicity of Essential Oils on Photosynthesis in *Mentha spicata* and *Thymus vulgaris*

Felicity Souza

ABSTRACT

Essential oils are chemicals extracted from medicinal and herbal plants, working as a defense system against microbes. Though essential oils are known to be able to chemically inhibit the growth of other plants, the extent to which they are toxic to their plant of origin hasn't been fully explored. This investigation aims to seek the potential negative effect of essential oils on the photosynthetic rate of plants containing the same oils. Essential oil compounds d-Limonene, linalool, and 1,8-cineole (independent variable) were applied to *Mentha spicata* (Common mint) and *Thymus vulgaris* (English Thyme) via a one μL droplet on a leaf. The dry mass of a leaf disc cut out from the droplet location was determined. The dry mass of a standard leaf disc (control) was also weighed, and the photosynthetic rate (dependent variable) was measured by calculating the percentage change in dry mass. A two-way ANOVA determined that the essential oils displayed a significant effect on the decreased dry mass in both species ($p < 0.001$). For example, linalool on mint had a mean percentage change of -47.9%, while linalool on thyme had a percentage change of -60.4%. This indicates the effects of autotoxicity since thyme, though mainly composed of linalool, exhibited a higher decrease in dry mass than mint, which lacks a high linalool content. These results indicate how leaf-chewing pests can transfer essential oils throughout the plant and enable auto-toxic effects that inhibit plant growth and germination. Autotoxicity could also be applied towards producing organic essential-oil based herbicides.

INTRODUCTION

This research study aims to investigate the potential autotoxicity of various essential oil compounds on two species of herbs. This knowledge can then be considered during the production of organic herbicides for agricultural use. Though chemical herbicides have been used in efficient control of weeds, they also contribute to air, soil, and water pollution, posing risks on human health. One of the most prevalent herbicides is glyphosate, a broad-spectrum herbicide that targets weeds by inhibiting an enzyme essential for plant growth. According to a 2019 analysis on glyphosate herbicide usage, almost 130 pounds were being

sprayed per square mile in the US (NBC, 2022). Nowadays, almost 60% of non-selective herbicides on the global market contain this active ingredient (Costas-Ferreira, Durán, Faro, 2022). However, contemporary agriculture now investigates reducing the use of these synthetic herbicides, and in turn seeing a development in essential oil-based herbicides. By using natural molecules present within plants, crop productivity is maintained without a negative environmental impact.

These essential oil compounds, such as terpenes or phenols, are produced and stored within the oil ducts and trichomes of aromatic species like herbs and spices. They are known to be allelopathic,

meaning that they can inhibit the growth and germination of other plant species. For example, the oil of lavender was examined and showed complete suppression of wheat and barley seed germination at concentrations of 30 and 90 μL (Zheljazkov, Jeliashova, Astatkie, 2021). In this way, use as an herbicide against noxious weeds is effective.

While allelopathy is well-researched, there is a minimal understanding of the concept of autotoxicity. Autotoxicity is an allelopathic phenomenon in which chemicals released into the environment by a singular plant species inhibits the growth of said species. This can be seen in the plant *Medicago sativa* (alfalfa), which releases allelochemicals that have lowered its own seed germination rates by up to 68.7% (Ghimire et al, 2021). These chemicals are made up of the same compounds found in essential oils, therefore indicating the potential autotoxicity of essential oil-producing plants.

In addition, the well-developed optimal defense theory suggests that in turn for a plant's defenses, resources are diverted away from growth and processes (Gershenzon, Ullah, 2022). Essential oils are a defense system that can prevent other species from germinating in close proximity. This indicates how plants that produce essential oils may be compromising the protection of vital cells and tissues. If these chemicals are leaked into locations other than the natural oil ducts present within a leaf, the plant may experience a decrease in productivity. Understanding the auto-toxic effects of essential oils could potentially help the development of improved organic herbicides.

Almost no studies have been done on this topic, with one notably focusing on the photosynthetic rates of wild parsnip, parsley, and rough lemon with compounds like γ -terpinene and α -pinene (Gog et al, 2005). This project aims to expand on this research by utilizing *Mentha spicata* (Common mint) and *Thymus vulgaris* (English thyme), as well as monoterpenes d-Limonene, linalool, and 1,8-cineole, which are major chemical components of essential oils. To evaluate toxicity, reductions in photosynthetic capacity were measured through the quantification of accumulated dry mass. It is hypothesized that if essential oils are applied to the mint and thyme leaves, then over time there will be a decrease in biomass of that leaf, indicating a declined process of photosynthesis. In addition to this, essential oils that are more abundant in mint than thyme will lead to a greater decrease in mint leaf biomass, and vice versa.

METHODOLOGY

One of each *Mentha spicata* and *Thymus vulgaris* plants were bought from Gray Barn Nursery and Garden Center, and soil type was assumed to be kept constant as a variable. During the experiment, plants were kept under the same plant light and given the same amount of water.

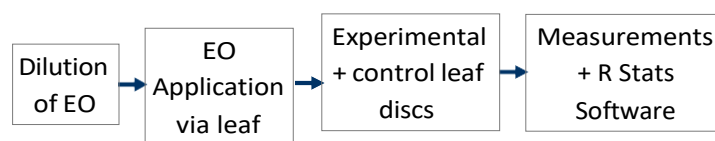


Figure 1. A simplified flowchart of the data collection procedure.

Before the topical application onto a leaf, the essential oils needed to be diluted from their 100% concentration. In a 5 mL vial, 1/8 teaspoon of canola oil was mixed with 3 drops of essential oil (d-limonene, linalool, 1,8-cineole) to create 3 diluted mixtures of each. Essential oil components were

bought online from Premier Research Labs. Canola oil acts as a carrier oil, which is a non-volatile component that works as a medium for dilution and application.

For the topical application, a micropipette was used to withdraw 1 μL of the diluted d-limonene essential oil (EO). Then, a 1 μL droplet of d-limonene was applied onto one side of a mint leaf and the droplet was punctured with an insect pin, piercing straight through the leaf to allow entry. After 2 hours had elapsed, a cork borer was used to extract a leaf disc from the area where the puncture was made. Near simultaneously, a second leaf disc of comparable size was cut out from the untreated half of the mint leaf, serving as a control to use in calculating the percent change in dry mass. These steps are visualized below in Figure 2. Then, both leaf discs were dried for a period of 10 hours. The dry mass in milligrams of the two leaf discs were measured using a balance scale and recorded in Table 1. With these two values, the percentage change in accumulated biomass over the 2-hour period was calculated and documented in Table 2.

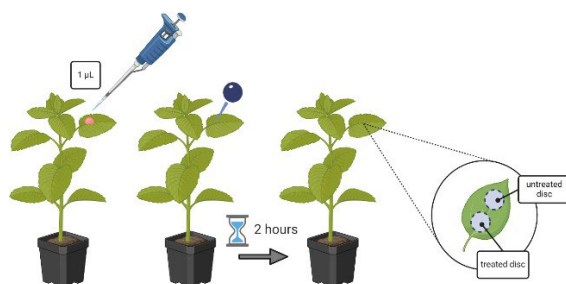


Figure 2. A visualization of procedure steps from EO application to cutting out leaf discs.

Note: Created by Felicity Souza on BioRender.com.

The above procedure was repeated using essential oils linalool and 1,8-cineole, as well as a control pinprick (without essential oils), conducting

each experiment on separate leaves. This was also repeated using English thyme leaves for comparative analysis. A total of 24 trials were completed, with 3 trials for each essential oil and plant combination.

The data was then compiled into an Excel spreadsheet, and a numeric summary was run using R Statistical Software to find a mean and median of the data set. Following, three different ANOVA tests (one-way, two-way, interaction) were run to find a p-value. A one-way ANOVA test is used when one categorical independent variable is tested on one quantitative dependent variable. The independent variable should have at least three categories; for this study, the EO variable encompasses d-limonene, linalool, and 1,8-cineole. A two-way test would look at the effects of two independent variables, such as EO and plant. An interaction two-way ANOVA test differs from the normal two-way because it evaluates the combined effect of both independent variables on the dependent variable.

The Akaike Information Criterion test was then run to find which ANOVA test was best representative of the data. A graph was made with ± 2 SEM and error bars to analyze the data, as shown in the results section.

Topical Application: Dry mass (mg) of Leaf Disc						
	Mint			Thyme		
	Tria 1 1	Tria 1 2	Tria 1 3	Tria 1 1	Tria 1 2	Tria 1 3
d-Limonene	17.7	17.7	17.7	14.3	14.3	14.3
	7	8.5	8	3	11	10.5
linalool	17.7	17.7	17.7	14.3	14.3	14.3
	7	8	9.5	3	4.5	6
1,8-cineole	17.7	17.7	17.7	14.3	14.3	14.3
	7	14.5	15	3	9	9
Control (pinprick only)	18	17	18	14	14	15
	18	17	18	14	14	15

Table 1. Data table with dry mass (mg) of leaf discs

Note: The top value in each box represents the control mass of the leaf disc before essential oil application. The bottom value represents the mass after a 2-hour period of essential oil application.

Topical Application: Percent Change (%) of Dry Mass of Leaf Disc						
	Mint			Thyme		
	Tria 1 1	Tria 1 2	Tria 1 3	Tria 1 1	Tria 1 2	Tria 1 3
d-Limonene	-60.4	-51.9	-54.8	-37.0	-23.0	-26.5
	5	8	0	6	8	7
Linalool	-43.5	-53.8	-46.3	-54.5	-68.5	-58.0
	0	0	3	5	3	4
1,8-cineole	-20.9	-18.0	-15.2	-44.0	-37.0	-37.0
	0	8	5	6	6	6
Control	0	0	0	0	0	0

Table 2. Data table with percent change in dry mass of leaf discs

RESULTS

From the numeric summary ran, it was determined that there was an overall negative percent change in dry mass for the data set, with a mean of -31.30%. Though only 3 trials were conducted for each essential oil and plant combination, the median, which was -37.06%, is close to the mean and shows that there are no significant outliers.

Following this, an interaction ANOVA model was ran, which evaluated the combined effect of the EO and plant variable on percent change in dry mass. In Figure 3, the EO variable and combined EO and plant variable are indicated to be both statistically significant at $p < 0.001$. However, the plant variable is not independently significant, and this is most likely due to the lack of variety in plant species since both mint and thyme are herbs. The residual sum of squares value shown is also relatively high at 354, indicating a potential deviation from the standard mean.

```
> interaction <- aov(dry.mass ~ eo*plant, data = plant.data)
> summary(interaction)
              Df Sum Sq Mean Sq F value    Pr(>F)    
eo              3   9773    3258  147.133 7.37e-12 ***
plant           1     18      18   0.824    0.378    
eo:plant        3   1978     659  29.783 8.79e-07 ***
Residuals     16     354       22                      
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 3. Interaction ANOVA test to determine statistical significance.
Note: Created by Felicity Souza on R Statistical Software.

The first hypothesis stated how over time, a decrease in biomass would be noticeable after essential oil application. This is supported by the graph displayed in Figure 4 because all experimental groups have a negative percent change in biomass. This means that for the period of 2 hours following essential oil application, one plant leaf produced less protein and organic matter in the treated disc

location than in an untreated location. The second hypothesis, which stated that essential oils more abundant in mint than thyme would display a greater decrease in leaf biomass, and vice versa, was also supported. This is because mint, which has a contains a high concentration of limonene in comparison to thyme, exhibited a greater decrease in dry mass at -55.7% instead of -28.9% for thyme. This same pattern can be applied towards experimental group 2, which tests for linalool.

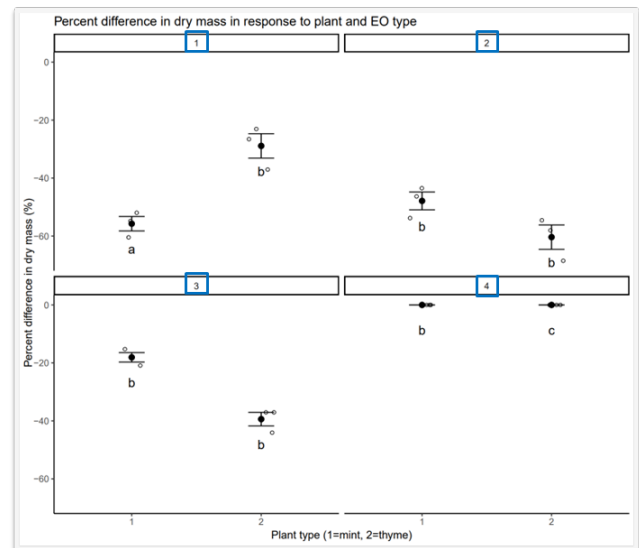


Figure 4. Percent difference in dry mass in response to plant and EO types

Note: Plotted graph with ± 2 SEM and error bars representing percent change in dry mass (y-axis) depending on plant type (x-axis) and 4 different essential oil groups (boxed in blue). The groups are limonene (1), linalool (2), cineole (3), and a control (4). Created by Felicity Souza on R Statistical Software.

CONCLUSIONS

The project was successful in its goal of identifying how essential oil compounds can inhibit their plant of origin by weakening the rate of photosynthesis in a leaf. It was displayed that over the two-hour period of essential oil application and absorption, both mint and thyme plants saw a lowered production in organic matter at the tested disc location. In an application of this knowledge to the natural world, pests that may chew on the leaves

of essential oil-producing plants can transfer the oils throughout the plant via the mandibles, causing additional damage to the plant's development. This research can also be considered when developing essential oil-based herbicides, which may be able to target weeds that have a higher concentration of a specific compound.

A supplemental study could involve expanding the range of compounds used within the project. This could include chemicals such as sesquiterpenes or phenols, which are less common but present within the constituents of essential oils. Monitoring nitrogen or phosphorus uptake would help gain a more holistic understanding of the plant's health in addition to photosynthesis. Measuring photosynthetic rate via a LI-COR system would test with a higher accuracy, compared to dry mass in this study. This portable device collected

real-time information by integrating the measurements of leaf gas exchange and chlorophyll A fluorescence into a complete assessment of photosynthesis.

Future research would also benefit from keeping plants in a more controlled laboratory environment. Despite keeping control of sunlight and water variables, the study made use of nursery-bought species. Managing the herbs throughout the entire growth process from germination would better handle discrepancies in specific growth stages; one leaf may be in a different stage than another, thus producing inconsistent amounts of biomass. Furthermore, the effect of insects on essential oil producing plant species could be investigated in a field study. This would provide a comprehensive understanding and application of autotoxicity in a real-world environment.

BIBLIOGRAPHY

- Costas-Ferreira, Carmen, et al. "Toxic Effects of Glyphosate on the Nervous System: A Systematic Review." *International Journal of Molecular Sciences*, vol. 23, no. 9, 21 Apr. 2022, p. 4605, www.ncbi.nlm.nih.gov/pmc/articles/PMC9101768/, <https://doi.org/10.3390/ijms23094605>.
- Gershenzon, J., & Ullah, C. (2022). Plants protect themselves from herbivores by optimizing the distribution of chemical defenses. *Proceedings of the National Academy of Sciences*, 119(4). <https://doi.org/10.1073/pnas.2120277119>
- Ghimire, B. K., Ghimire, B., Yu, C. Y., & Chung, I.-M. (2019). Allelopathic and Autotoxic Effects of *Medicago sativa*—Derived Allelochemicals. *Plants*, 8(7), 233. <https://doi.org/10.3390/plants8070233>
- Gog, L., Berenbaum, M. R., DeLucia, E. H., & Zangerl, A. R. (2005). Autotoxic effects of essential oils on photosynthesis in parsley, parsnip, and rough lemon. *Chemoecology*, 15(2), 115–119. <https://doi.org/10.1007/s00049-005-0294-8>
- Jefferies, D. (2022, October 10).
A potentially cancer-causing chemical is sprayed on much of America's farmland. Here is where it is used the most. NBC News. <https://www.nbcnews.com/data-graphics/toxic-herbicides-map-showing-high-use-state-rcna50052>
- Satyal, P., Murray, B., McFeeters, R., & Setzer, W. (2016). Essential Oil Characterization of *Thymus vulgaris* from Various Geographical Locations. *Foods*, 5(4), 70. <https://doi.org/10.3390/foods5040070>
- Snoussi, M., Noumi, E., Trabelsi, N., Flamini, G., Papetti, A., & De Feo, V. (2015). *Mentha spicata* Essential Oil: Chemical Composition, Antioxidant and Antibacterial Activities against Planktonic and Biofilm Cultures of *Vibrio* spp. Strains. *Molecules*, 20(8), 14402–14424. <https://doi.org/10.3390/molecules200814402>
- Zheljazkov, V. D., Jeliaskova, E. A., & Astatkie, T. (2021). Allelopathic Effects of Essential Oils on Seed Germination of Barley and Wheat. *Plants*, 10(12), 2728. <https://doi.org/10.3390/plants10122728>

Investigating the Success of Microbial Treatment for Gluten Digestion

Saahiti Mamidipaka

ABSTRACT

Grains such as wheat, rye, and barley are integral to the diet of several populations. However, it has recently been discovered that these gluten containing grains have been the cause of many hormonal and inflammatory disorders. This project aims to investigate the microbe that can best facilitate gluten digestion through an in-vitro simulation. *Bifidobacterium bifidum*, *Bifidobacterium longum*, or *Lactobacillus Acidophilus* microbes (Independent Variables) were first inoculated onto gluten-containing agar plates and incubated for 48 hours to simulate the amount of time food takes to digest. The growth of each bacterial colony was then observed to measure the cell growth due to gluten metabolism (dependent variable). The hypothesis was that the microbe *Bifidobacterium bifidum* would have the greatest change in growth. This was then proven correct with the values for colony count, total area covered by bacterial growth, and average percent area covered by each individual bacterium (146, 46045 mm², and 283.5 mm² respectively). The experiment's results also showed that any bacteria from the *Lactobacillus* class should not be used for probiotic treatment of gluten due to its growth being extremely similar to the growth of the control which showed that the bacteria were unable to properly grow on a gluten agar medium. Upon testing for the efficacy of *Bifidobacterium bifidum* compared to *Bifidobacterium longum* due to their similar growth patterns and statistics a t-test was conducted from which a p-value of 5.15×10^{-8} was calculated indicating that there is a significant growth observed on the *Bifidobacterium bifidum* plates.

INTRODUCTION

This study tests the use of *Bifidobacterium bifidum*, *Lactobacillus acidophilus*, and *Bifidobacterium Longum* in an in-vitro simulation of the human gastrointestinal system to determine which microbiota can best digest gluten. In the 21st century, terms such as gluten sensitivity, gluten-free, gluten intolerance, and celiac disease have gained prominence in the realm of dietary health. However, the impact on those who have these dietary restrictions is extreme and detrimental to their quality of life. In the United States, 30% of all Americans tend to avoid gluten (Sampedro, 2013). Of this 101 million, only approximately 2 million Americans have celiac disease and 20 million have gluten intolerance. Therefore, it is inferable that then

79 million Americans are on a gluten-free diet despite not being affected by any gluten – related disorder.

Gluten is a structural protein that is found in cereal grains, such as wheat, rye, and barley. Gluten itself is made up of two sub proteins: prolamin and glutenin. These are naturally occurring and trigger celiac disease patients' bodies. Gluten has viscoelastic and adhesive properties making bread and most baked products stretchy, chewy, and elastic. The formation of gluten occurs when glutenin molecules cross link and form disulfide bonds, resulting in a microscopic network attached to gliadin (a type of prolamin) contributing viscosity to the dough. Then, if this dough gets leavened with

yeast, the fermentation process produces carbon dioxide bubbles which get trapped by the microscopic network thereby causing the dough to rise. The baking process then coagulates the gluten and stabilizes the shape of the final product - in this case bread.

When gluten containing products need to be digested, it first gets chewed into small pieces, followed by swallowing with saliva, and then travels down the esophagus. Then it reaches the stomach and gets broken down into smaller pieces, the acids in the stomach also beginning to break down the food into protein form. This broken-down mixture gets sent to the duodenum, the beginning of the small intestine. Inside the small intestine are small hairlike structures called villi which absorb nutrients, vitamins, and minerals, but when gluten enters the small intestine it can trigger immune cells to defend the body. These immune cells are triggered due to conditions such as celiac disease or gluten allergy because the body is irritated by these foods and must protect against a harmful stimulus to prevent further irritation. In extreme cases it can even lead to death which is why the immune cells need to be on high alert. As a result of this defense mechanism, the body's immune cells destroy the villi and flatten them, so they cannot absorb the nutrients from the rest of the food consumed. This results in the body being deprived of vitamins, nutrients, and minerals which can lead to adverse side effects such as acne, diarrhea, bloating, malnutrition, depression, anxiety, etc. It is important to note that these side effects only occur in the population whose body is more sensitive to gluten and not in everyone, although gluten is still harmful

to most people's health despite them avoiding it or not.

A similar phenomenon happens to individuals who are lactose intolerant. However, in lactose intolerance it is an issue with the body not producing enough of the enzyme lactase to counteract the lactose in milk products. Lactase breaks down the lactose sugar into 2 more absorbable compounds: glucose and galactose. This becomes much easier to absorb and change into carbohydrates due to the body's familiarity with it. These carbohydrates are then used for energy. Now, if the lactase was not present (like in those with lactose intolerance) the common side effects of not being able to absorb this lactose are diarrhea, bloating, nausea, and belly pain. These symptoms have an overlap with the ones of gluten intolerance and celiac disease, particularly the bloating and diarrhea symptoms indicating a potential relationship in the treatment mechanisms of the two (Alkalay 2021).

When thinking about this problem from a different perspective, it can be determined that the microvilli can be regenerated when adhering to a strict gluten-free diet. However, the problem is that many foods can be contaminated with wheat products despite being considered gluten-free. A study conducted at Università La Sapienza di Roma's Public Health and Infectious Disease Center in Rome, Italy stated that using lactic acid bacteria to ameliorate gluten before consumption can hydrolyze the gliadin protein in gluten and as a result protect and support the regeneration of intestinal mucosa which contains microvilli (Cristofori et. al, 2020). This study discussed the feasibility of different

probiotic supplementation to hydrolyze gluten peptides for those with celiac disease. Hydrolyzing gluten peptides refers to the process of breaking down gluten peptides through the process of hydrolysis into smaller and more digestible chunks of protein. Later in the article it can be determined that the two most common probiotics are of the classes *Bifidobacterium* and *Lactobacillus*. These two classes of probiotics have been the most beneficial in hydrolyzing gluten proteins which was convincing evidence for this study to use probiotics from those two classes. *Lactobacillus* probiotics were the most beneficial however, due to their ability to create nitrogen from gluten peptides which assisted in carrying them across the cell membrane. In this study, however, the researchers used a mixture of *Lactobacillus alimentarius* 15M, *Lactobacillus brevis* 14G, *Lactobacillus sanfranciscensis* 7A, and *Lactobacillus hilgardii* 51B. That is where this article differs from this research project because a mixture of already successful proteins was used rather than testing new potential bacteria for probiotic treatment of gluten peptides. This research, led to further research into the types of lactic acid bacteria that can benefit the gut microbiota but haven't been used in testing yet. This makes it easier to standardize this class of bacteria to making gluten digestion more feasible for those with intolerances, allergies, or sensitivities.

In this investigation, the bacteria identified for their advantageous properties in the digestive system include *Bifidobacterium bifidum*, *Lactobacillus acidophilus*, and *Bifidobacterium Longum*. These probiotics were each chosen because of their current known contributions to the human

digestive system. *Bifidobacterium bifidum* is known for its abilities to break down carbohydrates which is useful due to gluten containing foods being primarily carbohydrates. *Bifidobacterium Longum* is known for its ability to break down fibrous substances which would be useful for foods containing whole wheat. *Lactobacillus Acidophilus* is known for promoting overall gut health and digestion. Notably, these bacterial strains have yet to undergo testing in the presence of gluten to assess their impact on the digestive system. The primary objective of this study is to examine the efficacy of *Bifidobacterium bifidum*, *Lactobacillus acidophilus*, and *Bifidobacterium Longum* within a simulated human gastrointestinal environment. This environment will be simulated in terms of temperature, fluids, and time.

This study hypothesizes that *Bifidobacterium bifidum* is the bacteria that will be the most effective in the simulated digestive process of gluten-related products due to its nature of being known to break down carbohydrates within the body.

MATERIALS AND METHODS

Materials

The materials utilized in the research serve specific purposes in simulating and analyzing the digestive process. Agar agar powder and distilled water are combined and sterilized to create gluten agar agar plates for bacterial growth. Wheat flour is added to nutrient agar for suitable bacterial cultivation. pH strips, along with sodium hydroxide or hydrochloric acid, ensure optimal pH conditions for bacterial growth. Sterile petri dishes and inoculating loops are used to culture and spread bacterial solutions. Incubation at 40°C in an incubator promotes

bacterial growth on labeled petri dishes. Simulated gastric fluid (SGF) and simulated intestinal fluid (SIF) mimic stomach and intestinal conditions.

Methods

The preparation of gluten agar agar plates involved mixing 6 teaspoons of agar agar powder in 3 cups of distilled water until dissolved, followed by sterilization of the mixture using an autoclave. Subsequently, the sterilized nutrient agar is combined with 2 teaspoons of wheat flour when it has cooled down to 120°C to create a total of 12 plates – 3 for each bacterium and 3 control plates. The pH of the nutrient agar was assessed using pH strips to ensure suitability for bacterial growth (within the range of 5.0 – 7.0). The pH in this range was necessary because that is the pH on which growth of these probiotic supplements is facilitated. Adjustment of pH, if necessary, was achieved by adding sodium hydroxide or hydrochloric acid. Before pouring the agar medium into the petri dishes, there needed to be a layer of simulated gastric fluid coating the bottom to replicate gastric conditions. The agar medium could then be poured into sterile petri dishes and be subjected to additional sterilization in the autoclave.

Following that process, the three probiotic treatment solutions were to be made. Combine 1.42 grams of the probiotic supplement powder with 100 mL of distilled water (or as per the dosing instructions for adults on the nutrition label). Use of a 10 µL micropipette and placement of this solution onto your starting point (indicated by the “x” in figure 2). Using inoculation loops of size 10 µL, the solution is spread onto 3 petri dishes for each bacterium (refer to figure 1 for bacteria streaking

techniques). The spreading of bacteria was not necessary for the control dish. Petri dishes were labeled with the bacteria’s name and date to ensure accurate data collection. Incubation of the labeled dishes occurred at 37°C for a variable period, up to a maximum of 48 hours, contingent on growth progress. This refers to the series of growth cycles that bacterium can go through before reaching its full growth potential.

How to Streak Bacteria onto a Petri Dish

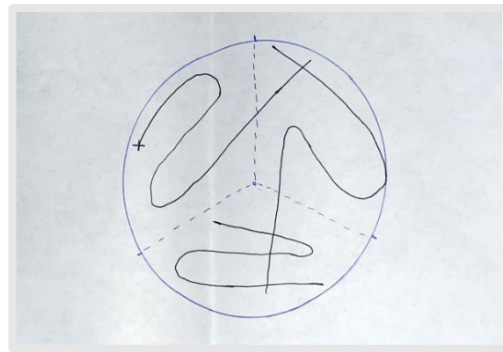
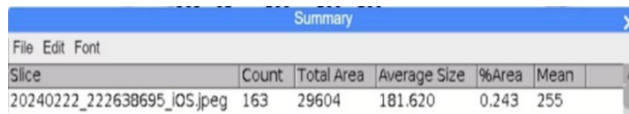


Figure 1 :Image of petri dishes with the pattern of the way the bacteria should be spread onto it with an inoculating loop. The “x” is indicative of the point where the micropipette solution should be placed and the starting point for the inoculation.

To analyze these growth patterns, photographs were taken of each of the bacterial plates and put into the ImageJ Analysis software to get adequate data. The image was first clipped in a circular pattern in order to avoid getting edges of the petri dish in the frame. Then the image was converted into a binary image and 16 bit to avoid color bias done by the software. The threshold of the bacterial growth was then highlighted for the analysis tool to later count. The image was then converted into watershed (to separate out different colonies) and sharpened to highlight the edges of each of the individual colonies. The bacterial growth was then analyzed using the ImageJ analysis tool to get the measurement of colony count, total area (mm²), and

individual area(mm²). Refer to Figure 2 for an example summary of results that ImageJ provides.

ImageJ Results Summary



Slice	Count	Total Area	Average Size	%Area	Mean
20240222_222638695_OS.jpeg	163	29604	181.620	0.243	255

Figure 2: Image of the summary statistics given in ImageJ. Colony count is given by counting all the colonies provided, counting the total area (mm²) of colonial growth, average size (mm²) by averaging total area, and percent area by averaging the percent covered by each colony. Mean stayed standard as 255 for all, so it was not utilized in the data collection process.

DATA TABLES

Data Table 1.1(Control)

	Plate 1	Plate 2	Plate 3	Average
Colony Count	49.67	0.33	0	16.67
Total Area	14312.67	21.33	0	4778.0
Avg Colony Area	212.132	21.33	0	77.88

Data Table 1.2 (Bifidobacterium longum)

	Plate 1	Plate 2	Plate 3	Average
Colony Count	93	60.67	12.67	54.44
Total Area	28162	13760.67	1888	14603.56
Avg Colony Area	386.90	217.675	298.88	301.15

Data Table 1.3(Lactobacillus acidophilus)

	Plate 1	Plate 2	Plate 3	Average
Colony Count	2	0	2.67	1.56
Total Area	1234	0	1265	833
Avg Colony Area	450.61	0	396.39	282.33

Data Table 1.4 (Bifidobacterium bifidum)

	Plate 1	Plate 2	Plate 3	Average
Colony Count	241.33	9.67	51.33	146.33
Total Area	74500.67	1717.33	17589.33	46045
Avg Colony Area	282.65	162.21	284.32	283.49

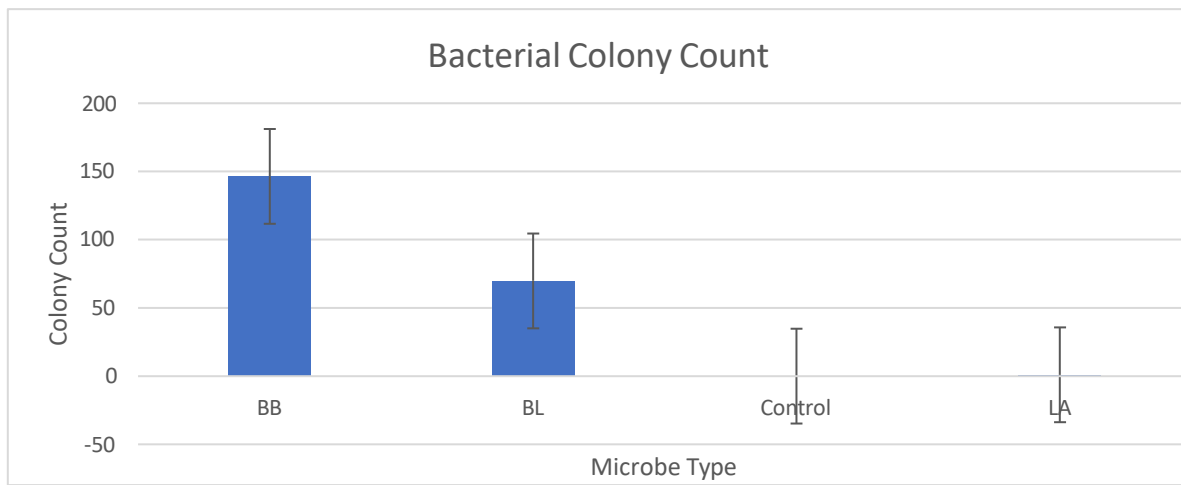


Figure 3 Data derived from ImageJ analysis of the colony count between the 3 plates probiotic growth. *Bifidobacterium bifidum* had the largest growth as expected.

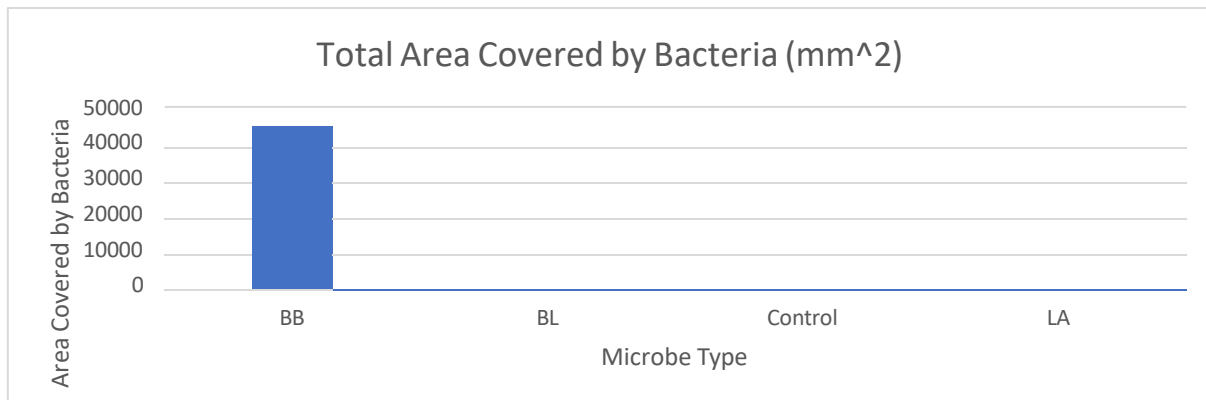


Figure 4

Figure 4: The total area covered by each probiotic strain's growth. Data for these graphs was received from ImageJ. *Bifidobacterium bifidum* has the largest growth as expected.

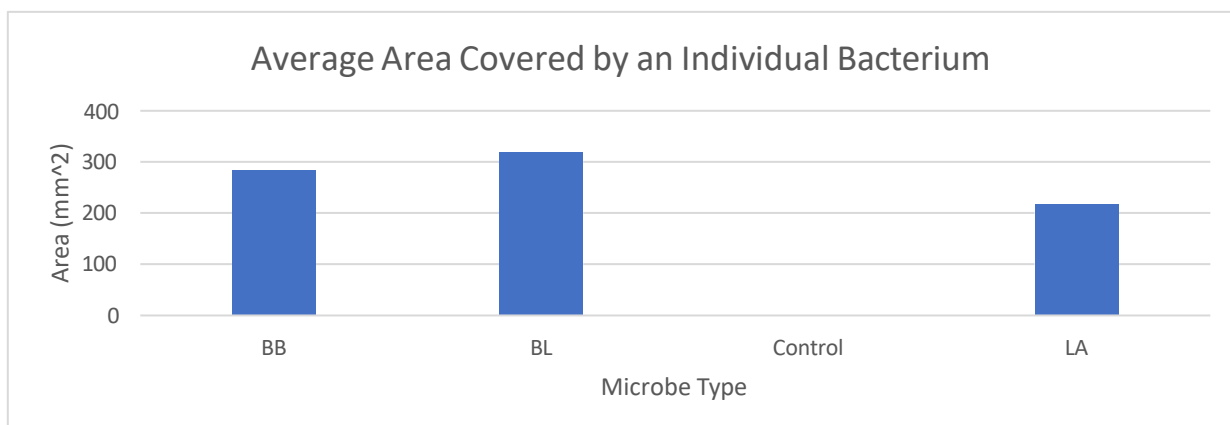


Figure 5 The average individual area covered by each type of strain of bacterial growth. Data for these graphs was received from ImageJ. *Bifidobacterium longum* has the largest growth, which is unexpected.

RESULTS

The hypothesis that *Bifidobacterium bifidum* would have the most growth in terms of total area covered by all bacterial colonies (mm²) as well as mean bacterial colony count was correct. The total area covered by *Bifidobacterium bifidum* is 46,045 mm² which may be due to an outlier on *Bifidobacterium bifidum* plate 1. The whole plate was covered with a white wafer layer of bacteria. There were approximately 146 colonies growing on the *Bifidobacterium bifidum* plates.

A paired t-test to further solidify that *Bifidobacterium bifidum* was the most effective in comparison to *Bifidobacterium longum* due to its bacterial growth. The null hypothesis was that in comparison to *Bifidobacterium longum*, the growth of *Bifidobacterium bifidum* was insignificant. The p-value was 5.15×10^{-8} , which is less than 0.05

mechanism which doesn't have any drawbacks and promotes healthy gut bacteria growth which can benefit all individuals and can potentially solve problems differing from just gluten-related disorders.

DISCUSSION

Due to the adverse effects of gluten-related disorders upon gluten consumption for many individuals, doing an in-vivo test with these probiotics would not be recommended especially for a high school student to conduct. More in-vitro testing with tissue cultures must be conducted with these probiotics to test the efficacy of these probiotics in the digestive system before this can be applied to those with fatal implications upon consumption of gluten. This study was also conducted in a high school lab without industry standard equipment which could affect the

indicating statistical significance. This means that the null hypothesis was rejected, and it was concluded that there was convincing evidence that *Bifidobacterium bifidum* was the most effective probiotic supplement in aiding in a simulation of the digestive system to test gluten digestion.

CONCLUSIONS

Due to the statistically significant data, we can conclude that *Bifidobacterium bifidum* could be the most applicable as a probiotic supplement to consume before gluten consumption because it had statistically significant data to be the most successful probiotic.

This is important to the future of research on gluten-related disorders because it sets a precedent for the future of treatment to a disorder which currently only has one treatment: a gluten-free diet. The use of probiotic supplementation is a treatment potential of probiotic growth which needs to be investigated further before applying this to any medical scenarios of those with gluten-related disorders.

FUTURE STEPS

In the future, this experiment should be conducted in an industry standard laboratory with more controlled conditions. This is to ensure that no confounding variables affect the data that is collected. This studies results should also be applied to an in-vivo testing scenario to test if the probiotic supplementation occurs similarly inside the human body compared to the in-vitro simulation. This in-vivo testing should be conducted as a clinical trial with supervision from medical professionals to ensure adverse reactions do not occur for those with gluten-related disorders.

ACKNOWLEDGEMENTS

I'd like to thank my mentor Mr. James Boyd for assisting me with the setup of my lab procedures for this experiment. I'd like to thank Mrs. Kate Allender for her words of encouragement and guidance throughout this research. I'd like to thank my parents for their support through all my endeavors. Finally, I'd like to thank my younger sister for providing advice throughout this project and all the hardships experienced.

REFERENCES

- Alkalay M. J. (2021). Nutrition in Patients with Lactose Malabsorption, Celiac Disease, and Related Disorders. *Nutrients*, 14(1), 2. <https://doi.org/10.3390/nu14010002>
- Cristofori, F., Francavilla, R., Capobianco, D., Dargenio, V. N., Filardo, S., & Mastromarino, P. (2020). Bacterial-Based Strategies to Hydrolyze Gluten Peptides and Protect Intestinal Mucosa. *Frontiers in immunology*, 11, 567801. <https://doi.org/10.3389/fimmu.2020.567801>
- Diez-Sampedro, A., Olenick, M., Maltseva, T., & Flowers, M. (2019). A Gluten-Free Diet, Not an Appropriate Choice without a Medical Diagnosis. *Journal of Nutrition and Metabolism*, 2019.
- Gabler, A. M., & Scherf, K. A. (2020). Comparative Characterization of Gluten and Hydrolyzed Wheat Proteins. *Biomolecules*, 10(9), 1227.
- Roxas M. (2008). The role of enzyme supplementation in digestive disorders. *Alternative medicine review: a journal of clinical therapeutic*, 13(4), 307–314.
- Ali, B., & Khan, A. R. (2022). Efficacy of Probiotics in Management of Celiac Disease. *Cureus*, 14(2), e22031. <https://doi.org/10.7759/cureus.22031>
- Ramedani, N., Sharifan, A., Gholam-Mostafaei, F. S., Rostami-Nejad, M., Yadegar, A., & Ehsani-Ardakani, M. J. (2020). The potentials of probiotics on gluten hydrolysis; a review study. *Gastroenterology and hepatology from bed to bench*, 13(Suppl1), S1–S7.

Supplementing Immunity: Investigating the Effect of Vitamin D Supplementation on Human Cancer T-cell Proliferation and IL-2 Secretion

Pavitra Dave

ABSTRACT

In the realm of immune-related research on natural products to protect against growing public health concerns, the influence of market-available vitamin D supplements, with lipid additives, on cellular behavior, remains a mystery. This project looks at the relationship between supplemental vitamin D serum with lipid ingredients, and cell death of human cancerous T cells (type of white blood cell). This study aims to unveil potential modulatory influences of this vitamin D, with lipid additive, on cellular responses, contributing crucial insights to the broader understanding of supplements' public health implications. The effect of varying concentrations of vitamin D (independent variable) is measured on the stimulated cells by measuring the resulting optical density values (dependent variable), which represents the viability of the cells. Cancerous T cells stimulated with certain compounds (PMA and Ionomycin) secrete a protein, IL-2, representing T cell growth. To further assess cell viability, this protein's secretion is measured in wells with given optical densities, and interpolation with a standard curve. It is hypothesized that the presence of varying vitamin D, and therefore lipid additive, concentrations would significantly affect the T cells' viability. Specifically, a potential linear relationship, with higher concentrations resulting in higher cell viability and more IL-2 secretion. Post-experimentation, ANOVA analysis for concentration-specific analysis of cell viability returned a significant $p = 0.000000018$, concluding that the concentrations of vitamin D used as experimental conditions significantly influenced cellular response. Cell viability analyzed in correlation with IL-2 secretion, a concentration dependent inhibitory effect on PMA+IO induced IL-2 secretion is found.

KEY TERMS

1. **T cell**- subtype of a white blood cell, part of the body's immune system.
2. **Interleukin-2 (IL-2)**- a type of cytokine, signaling molecule, and protein part of the body's immune system.
3. **Vitamin D**- A fat soluble vitamin essential for healthy bones, immune system, and other cellular functions. It can be obtained from certain foods, supplements, or sunlight.

INTRODUCTION

The purpose of this project is to systematically explore the response of cancerous T cells to varying concentrations of a commercially available vitamin D supplement formulated with an olive oil lipid carrier. Specifically, aiming to investigate how the supplement formulation influences both cellular

viability and interleukin-2 (IL-2) secretion: crucial aspects of T cell function and immune regulation.

The pervasive issue of vitamin D deficiency, affecting approximately 42% of the U.S. population, underscores a significant concern for public health (Forrest, Stuhldreher, 2011). Our skin's ability to synthesize vitamin D is tied to the wavelength of UV-B rays immitted from the sun;

lack of sun resulting in improper vitamin D levels. Specifically, diminished exposure to sunlight and the altered solar angle during, for example, the winter, significantly hampers the body's capacity to produce sufficient vitamin D. Consequently, research continues to unveil negative correlations between low vitamin D levels and various health issues, from causing an imbalance in calcium/phosphorus regulation, to bone and metabolic health (Holick, Chen, 2008). Sufficient levels of vitamin D have been further associated with decreased risks of cancers, including colorectal and bladder cancer (McCullough, et al., 2019). As awareness of these health implications has grown, the market has responded with the development of vitamin D supplements designed to address deficiencies and promote overall well-being.

Exploring the variety of vitamin D products available, reveals that these products exhibit variability not only in their vitamin D content, but also their formulations. A deeper investigation into the effects of the formulation contents used in the supplements is necessary to determine the influence of the constituents on cellular systems. Notably, the various carriers (which protect, stabilize, and support consistent delivery) play a pivotal role in influencing the bioavailability and efficacy of these supplements. In essence, creating more effective supplements.

Lipid carriers, in particular, are used to increase absorption and the overall potential effect of vitamin D. While there is evidence stating that lipids may contribute to immune modulating

effects themselves, such as the secretion of a set of immune proteins in mammals given olive oil diets, the effects in combination with vitamin D are unknown (Puertollano, et. al., 2007).

Building on this understanding of carriers, this research project aims to explore the effect of varying concentrations of a commercially available vitamin D supplement formulated with one of these lipid additives, olive oil, in JURKAT cells. Serving as the model system, JURKAT cells are a cancerous transformed human T lymphocyte cell line, originally derived from a 14-year-old boy with T cell leukemia. While a strong correlation between vitamin D and cancer risk has been reported, the specific effects of vitamin D on human T cells remain relatively unclear. To quantitatively explore the potential effect, the cell's secretion of a protein will be used. Known for its upregulation of proliferation and activity of T lymphocytes (T cells), IL-2 is a crucial cytokine, or signaling protein, in the immune system. Additionally, cellular proliferation will be assessed to evaluate the overall impact of the supplement on cell growth and activity. By simultaneously examining both IL-2 secretion and cellular proliferation, this study aims to uncover the nuanced relationship between varying concentrations of the vitamin D supplement and their effects on JURKAT cells, thereby examining the potential mechanisms underlying immune modulation and cancer cell behavior. This could guide the optimization of strategies for supplementation of vitamin D deficient patients and individuals at a higher cancer risk.

It is hypothesized that the incorporation of varying concentrations of a commercially available vitamin D supplement formulated with olive oil as a lipid additive will elicit distinct modulatory effects on cell proliferation as well as interleukin-2 (IL-2) secretion. Specifically, regarding cell proliferation, varying concentrations of the vitamin D supplement will likely affect cells differently, hypothesizing that higher concentrations will result in higher cell viability. It is anticipated that increasing concentrations of the supplement will lead to a dose-dependent increase in cell viability, reflecting the potential stimulatory effects of vitamin D on cellular proliferation. Evidence suggests that lipid additives influence the effectiveness and uptake of vitamin D. Therefore, the hypothesis could be formulated that this lipid specific market-available supplement may play a distinct impact on IL-2 secretion in the cells with each varying concentration.

METHODOLOGIES

Broken up into two major parts: two experiments (XTT cell proliferation assay, and ELISA IL-2 measurement assay)

Cell Culture preparation

JURKAT cells originally from the American Type Culture Collection must be removed from nitrogen tank storage.

To begin with culture preparation, thaw JURKAT cells in 37°C water bath. Transfer thawed cells into warmed RPMI40 media with 10% FBS and 2MM L-Gln. Centrifuge cells at speed 750 rpm for five minutes. Decant cells and

resuspend cells in warmed RPMI media. Over the next few days, the cells are counted under a microscope using a hemocytometer (see Figure 1a and 1b below) and trypan blue dyeing technique to assess viability and cell concentration.

For intended cell viability as well as protein secretion assays, the cell line JURKAT E6.1 must be split to the concentration of 1×10^5 c/mL, as well as 4×10^5 c/mL (two XTT, or cell viability, assay plates are created to compare/contrast noise readings of data).



Figure 1a: Hemocytometer used for cell counting.

:

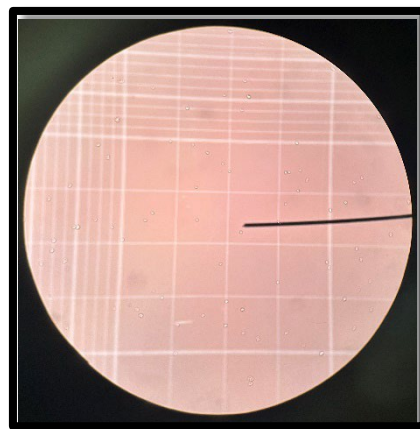


Figure 1b One 4X4 grid corner on hemocytometer, with cells, under microscope.

METHODOLOGY (Part 1: XTT Assay- Measuring Cell Viability Post Vitamin D

treatment): See Figure

For this experiment measuring the effect of the varying concentrations of vitamin D supplement on cell viability was achieved through an XTT assay. An XTT assay is a colorimetric assay used to quantify JURKAT cell viability based on the metabolic activity of the cells which is hypothesized to alter based on cell treatment which is varying vitamin D concentrations.

This assay requires the plating of the cells with phorbol 12-myristate 13-acetate (PMA) and Ionomycin- compounds which induce the production of cytokines (IL-2) and other immune response mediators. Optimization of PMA is essential to ensure the conditions used in the assay are optimal for stimulating T cell activation and subsequent cytokine secretion. Therefore, the XTT method begins with the standardization of PMA, using six different compound concentrations for testing and the creation of an optimization curve. Begin by splitting JURKAT cells to concentration required for this optimization test- 1×10^5 c/mL. Prepare six PMA dilution calculations. The concentrations of PMA were 0.1, 1, 2.5, 5, 10, and 20 ng/mL. See Table 1 below for PMA optimization results, and Figure 2 for PMA optimization data visualization.

Table 1: Average Absorbance (Optical Density 475-660) from PMA compound standardization.

Tx Conc	Absorbance (OD475-660)	Std Dev
Media	0.019	0.002
Cells	0.641	0.082
Iono	0.784	0.106
C6	0.518	0.049
C5	0.502	0.040
C4	0.509	0.068
C3	0.501	0.058
C2	0.524	0.061
C1	0.227	0.040

Note: Higher Absorbance (OD475-660) indicates more cell activity, or more cells being alive

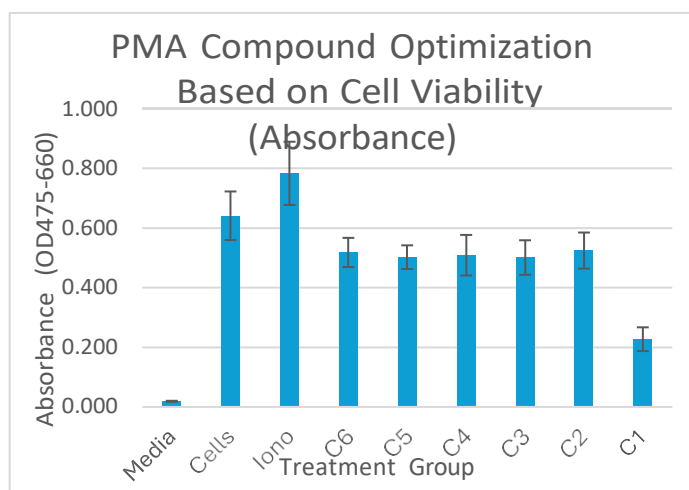


Figure 2: Visualization of PMA optimization data. Media, Cells, and Iono being controls to ensure experimental plate is error free. C6 (0.1 ng/mL PMA), C5 (1 ng/mL PMA), C4 (2.5 ug/mL), C3 (5 ug/mL PMA), C2 (10 ug/mL PMA), C1 (20 ug/mL PMA).

C2 (10ug/mL PMA) chosen as optimal PMA concentration for required XTT experiment, as can be seen by the sharp drop in Absorbance (OD475-660) value meaning PMA above the concentration of C2 has a cell death inducing effect.

Next, a vitamin D supplement concentrations range is formulated. Using 1 ug/mL as the dosage, a log scale with a few interval points is created to be the 5 different concentrations being tested on JURKAT T cells. With each well in the plate set to have 200 ug/mL, the divisions of cells, PMA/IO, and treatment are divided up. In each well, 100 uL of cells (at 1×10^5), 25 uL of PMA, 25 uL of Ionamycin, and 50 uL of treatment (vitamin D supplement at specific concentration).

To prepare serial dilutions of vitamin D3, PCR tubes or conical tubes are labeled sequentially from P1 to P5. See Table 2 below to note each label with its designated concentration. In the first step, the purchased vitamin D3 is diluted to create P1 (the first, most concentrated variation). Assuming the initial concentration of the vitamin serum is 990 $\mu\text{g/mL}$ as indicated on the vitamin bottle, 808.08 μL of the vitamin D3 solution is extracted. Next, P2 is prepared by combining 1000 μL of P1 with 1000 μL of media. Similarly, P3 is created by mixing 200 μL of P1 with 1800 μL of media. P4 is then obtained by adding 20 μL of P1 to 1980 μL of media, while P5 is prepared by combining 2 μL of P1 with 1998 μL of media. These serial dilutions allow for the generation of a range of concentrations of the

vitamin D3 solution for subsequent experimental use.

Note: the creation of these vitamin D concentrations requires an additional step of vortexing inside the hood. A decision that was made based on the oil separation of the serum due to its oil consistency and ingredient.

Table 2: Conical Title and Corresponding Vitamin D Supplement Concentration

Concentration Title	P1	P2	P3	P4	P5
Vitamin D Concentration (ug/mL)	100	50	10	1	0.1

The final experimental groups to be plated are the control groups. To ensure comprehensive evaluation, several controls were included in the plate setup. These included a vitamin D3 concentration without PMA and IO, serving to assess the vitamin D3-specific effects. Additionally, a positive control with MG132 confirmed external cellular responsiveness to stimuli. Controls with PMA and IO validated T cell activation, while IO alone assessed its isolated effects. Finally, control with cells and media alone accounted for nonspecific effects. These controls collectively confirm the reliability and validity of the experimental results.

With all treatment and control solutions prepared, plate them according to figure 3 below. Accounting for six replicates of each of the 5 vitamin D concentrations, and six replicates of each of the 5 controls. In this, practicing border well exclusion will minimize edge effects such as

evaporation or inconsistent temperature distribution.

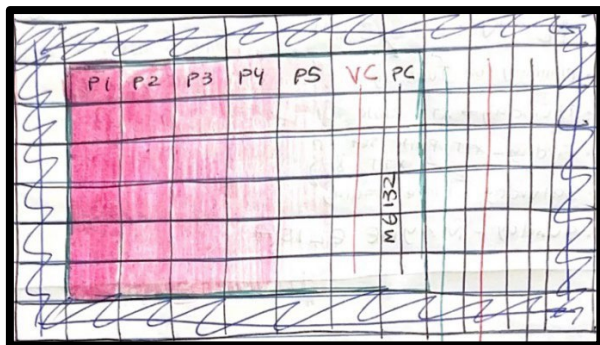


Figure 3:: XTT Assay 96-well Plate Set-Up: Vitamin D concentrations arranged on left of plate, highest concentration (P1) at left, and lowest concentration (P5) at right. Complete right hand side of plate (VC, PC, and last three columns) are the controls to compare cell viability and IL-2 reading against.

Stimulation of the cells with these serum and compound varieties requires multi-channel pipettes, as well as troughs to pour each dilution or liquid into for plating.

This experiment was ran with two plates, one with cells at 1×10^5 c/w and a second with cells at 4×10^5 c/w. Quantities of supplements as well as controls are consistent throughout both plates.

Prior to the plates being read, liquid supernatants from the cell plates are removed and re-plated for the next ELISA. Plates of supernatants (total of two, each from plates with the two different cell concentrations) are kept in fridge till ELISA assay.

For the plates with remaining cells, set up plate reading template on SoftMax_Pro for them to be read. Absorbance is read off, OD475–OD650, using a microplate reader. Final quantitative values are given in Optical Density and analyzed thoroughly using software such as

Excel. Higher optical density value correlates to more metabolic activity in the cells, essentially more cells being alive.

METHODOLOGY (Part 2: ELISA Assay)

ELISA (enzyme-linked immunosorbent assay) was conducted to quantify interleukin-2 (IL-2) secretion in JURKAT cells. ELISA kit from bio-technie was utilized in specific scenario, and kit instructions were followed. First, create buffer between 4.7-4.9 using pH strips. Coat plate with coating buffer day before performing ELISA assay.

Samples of supernatants are then added and incubated (after coating buffer) to allow IL-2 to bind to plate. After washing and blotting, detection antibody specific to IL-2 is added, followed by a secondary antibody. To induce color change for plate reading, substrate solution is added. Color change is measured on microplate reader. The resulting absorbance values used to determine IL-2 concentrations in the samples, interpolated from a standard curve generated using known IL-2 samples. This allows for the quantification of IL-2 secretion in response to varying concentrations of vitamin D supplementation in the cells.

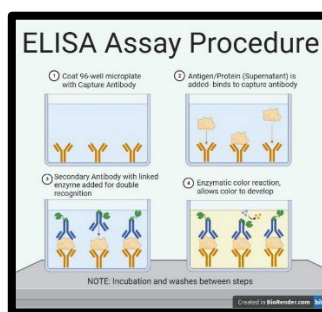


Figure 5: ELISA Assay Procedure: Procedure for testing IL-2 secretion quantitatively using ELISA assay. Step (1), coating microplate with antibody for capturing antigen. Step (2) displaying the capture of antigen due to recognition site. Step (3) introducing secondary antibody for antigen to be recognized by two domains. Step (4) adding substrate to solution to test color variance.

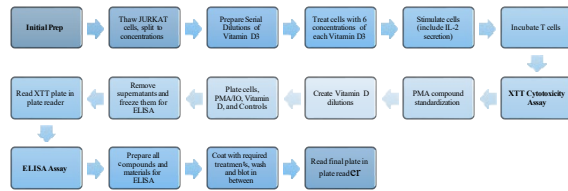


Figure 5: Whole procedure generalized
Overview of total procedure. Beginning with preparation of cells, to preparation of vitamin D3 for testing, onto treatment, stimulation, incubation, and analysis using XTT and ELISA assays.

DATA ANALYSIS

Data analysis was conducted using Microsoft Excel, where graphical representations were generated for both the XTT assay, as well as the ELISA experiment. Excel's graphing functionalities were used to visualize trends and patterns in the data. Gaining insight on how the treatment groups (which were a range of concentrations) affected the JURKAT T cells differently. Additionally, a statistical analysis was performed using analysis of variance (ANOVA) to assess the significance of difference observed among the experimental groups. This would help confirm that the vitamin D varying in concentration did in fact play a specific role on the cells (dependent to that concentration, or not). The ANOVAs on both the XTT as well as the ELISA yielded statistically significant results, indicating that the observed variations in cell viability (XTT) and interleukin-2

Source of Variation	SS	df	F	P-value	F crit
Between Groups	0.569677472	5	18.83956	1.86E-08	2.533555
Within Groups	0.181430167	30			

(IL-2) expression levels (ELISA) among different treatment groups were not due to chance.

Source of Variation	SS	df	MS	P-value	F crit
Between Groups	1691637.5	5	338327	1.93804E-29	2.53355
Within Groups	16061.4862	30	535.383		

RESULTS

In analyzing results, it is noted that there was no linear relationship between the concentration of the vitamin D3 supplement and cell proliferation. See Figure 6 for a clear visual of the relationship. Nevertheless, a one-way analysis of variance (ANOVA) revealed statistically significant differences in the effects of the various concentrations on cell viability. This suggests that although the relationship between concentration and cell proliferation may not follow a linear trend, the concentrations of the vitamin D3 supplement exerted distinct and measurable effects on cell viability. Further exploration into the observed non-linear relationship may be necessary to better understand the dose-response relationship. These findings highlight the importance of considering non-linear effects in dose-response studies and emphasize the need for comprehensive data analysis techniques to derive meaningful insights from experimental results.

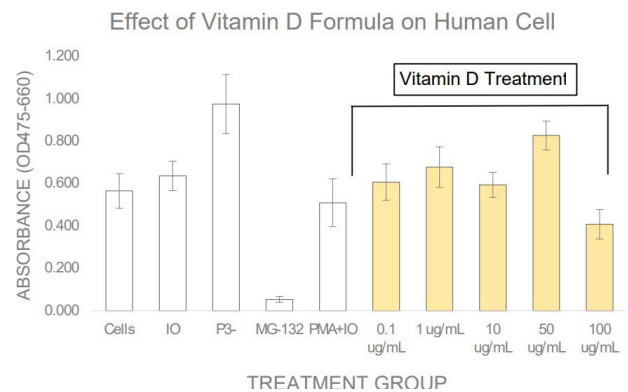


Figure 6: XTT Assay, Cell Proliferation Data Visualized
Data results from XTT Assay visualized. P0→P5 are 6 concentrations of vitamin D treatment. P3- is concentration P3 plus cells and media (no IO/PMA). MG-132 is positive control. PMA + IO stimulate IL-2 secretion in cells

Notably, examination of the data points on Figure 7 (scatter plot), reveal that data points were fairly close together, indicating consistency and reproducibility in the experiment results across the plate. This further strengthens the reliability of the findings and suggests that the results are replicable.

Figure 7: Scatter Plot of Absorbance Data Points in XTT Assay (Distribution Visualization)

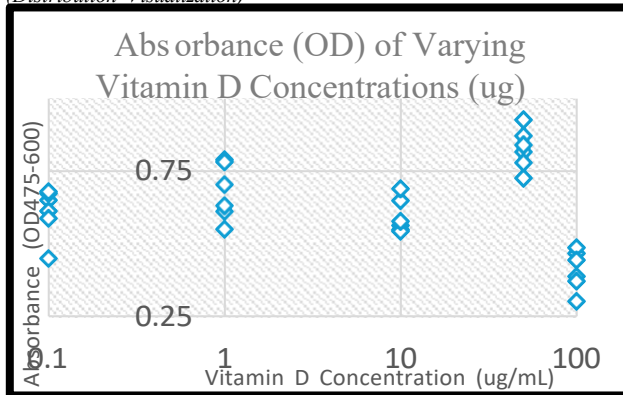


Figure 7: Distribution of Absorbance readings in plate- pointing out any potential outliers (limited).

ELISA results support as well as adds to results observed in the XTT cell proliferation assay. After the created ELISA standard curve was used to interpolate protein secreted by treatment groups, it is clear that the ELISA assay demonstrated a concentration-dependent inhibitory effect on PMA + IO-induced IL-2 secretion. This can be visualized in Figure 9. This observation suggests a correlation between cell death and low IL-2 secretion levels, indicating a potential mechanism through which vitamin D supplementation modulates immune responses.

This concentration dependent effect could potentially be due to the killing of cells with a higher vitamin D concentration (in reference to XTT results), however analysis of concentration 50 ug/mL shows that with a relatively high cell viability

count, IL-2 concentrations are still low. Going against the suggestion that cell death is the cause of the IL-2 protein secretion results.

Figure 8: Elisa Standard Curve

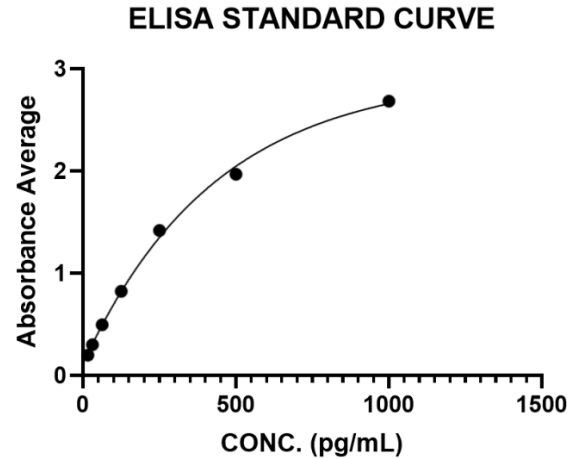
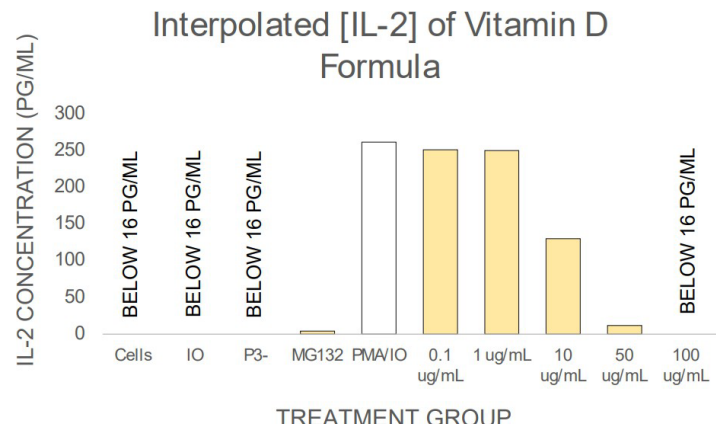


Figure 8 Standard curve generated for IL-2 concentration measurements. Graph depicts the relationship between known IL-2 concentrations and their corresponding absorbance values, obtained through ELISA. Standard curve will be used to interpolate and accurately quantify IL-2 values in treatment groups (visualized in Figure 9).

Figure 9: Interpolated IL-2 Concentrations in Unknown/Treatment Groups



ELISA graph illustrating interpolated IL-2 values. As the concentration of treatment (vitamin D) increases (visualized in the orange bars), there is a corresponding decrease in IL-2 concentration.

CONCLUSION

The absence of a linear relationship between concentration and cell proliferation in the XTT assay results highlights the complexity of the response of JURKAT cells to varying concentrations of the vitamin D3 supplement. Despite this, the significant

differences observed among the concentrations emphasize the potency of the supplement in modulating cell viability. The proximity of data points on the graph confirms the consistency and reproducibility of the experimental results, enhancing the reliability of the findings.

The non-linear relationship observed suggests that the effects of the vitamin D3 supplement on cell viability may be influenced by factors beyond simple concentration-dependent mechanisms. Further investigation into the underlying mechanisms driving the observed effects, such as potential saturation points, may provide valuable insights into the dose-response relationship. ELISA assay results, measuring IL-2 protein

secretion, validate and provide greater insight into the vitamin D's effect on the cells.

Additionally, XTT and ELISA results align with existing evidence suggesting vitamin D's inhibition of the NFkB pathway. This sheds light on the potential therapeutic implications of vitamin D supplementation in modulating cellular responses, particularly in the context of cancer treatment. In this study's case, while the observed death of cancer cells may be considered beneficial, it also highlights the need to further understand the mechanisms underlying vitamin D's effects as a supplement, on cellular pathways.

BIBLIOGRAPHY

- (2023). Global Vitamin D Market Insights Forecasts to 2030. *Spherical Insights*.
<https://www.sphericalinsights.com/reports/vitamin-d-market>.
- Bowles, J. (2021). Overwhelming Proof that Vitamin D3 Deficiency Causes Most Human Diseases. <https://jeffbowles.com/vitamin-d3-deficiency-causes-most-human-disease/>.
- Budhathoki, S., et. al. (2018). Plasma 25-hydroxyvitamin D concentration and subsequent risk of total and site specific cancers in Japanese population: large case-cohort study within Japan Public Health Center-based Prospective Study cohort. *The BMJ*. [10.1136/bmj.k671](https://doi.org/10.1136/bmj.k671).
- Forrest, K., and Stuhldreher, W. (2011). Prevalence and correlates of vitamin D deficiency in US adults. *Nutrition Research*. 31(1). <https://doi.org/10.1016/j.nutres.2010.12.001>.
- Holick, M., and Chen, T. (2008). Vitamin D Deficiency: A worldwide problem with health consequences. *The American Journal of Clinical Nutrition*. 87(4). <https://doi.org/10.1093/ajcn/87.4.1080S>.
- McCourt, A., Mulrooney, S., O'Neil, G., O'Riordan, E., and O'Sullivan, A. (2023). Serum 25-hydroxyvitamin D response to vitamin D supplementation using different lipid delivery systems in middle-aged and older adults: a randomised controlled trial. *British Journal of Nutrition*. <https://doi.org/10.1017/S0007114523000636>.
- McCullough, M., et al. (2019). Circulating Vitamin D and Colorectal Cancer Risk: An International Pooling Project of 17 Cohorts. *Journal of the National Cancer Institute*. [Circulating Vitamin D and Colorectal Cancer Risk: An International Pooling Project of 17 Cohorts - PMC \(nih.gov\)](https://doi.org/10.1093/jnci/kjz111)
- Moyer, V. (2014). Vitamin, Mineral, and Multivitamin Supplements for the Primary Prevention of Cardiovascular Disease and Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine*. <https://doi.org/10.7326/M14-0198>.
- Puertollano, M., et al. (2007). Significance of Olive Oil in the Host Immune Resistance to Infection. *Cambridge University Press*. [Significance of olive oil in the host immune resistance to infection | British Journal of Nutrition | Cambridge Core](https://doi.org/10.1017/S0007114507000636)
- Sasagawa, M., Cech, N., Gray, D., Elmer, G., Wenner, C. (2006). Echinacea alkylamides inhibit interleukin-2 production by Jurkat T cells. *International Immunopharmacology*. doi:10.1016/j.intimp.2006.02.003.

Surviving Salt: The Osmoprotective Effects of Exogenous Trehalose and Sucrose on Mung Bean (*Vigna radiata*) Growth Under NaCl Stress

Shuhui Yang

ABSTRACT

Millions of people worldwide have lost their farmland to soil salinization caused by sea level rise, saltwater intrusion, drought, and poor irrigation practices. Most crops like mung beans are salt sensitive which creates a need to develop salt-tolerant cultivars to feed a rapidly growing population. Carbohydrate osmoprotectants are natural plant sugars that defend against osmotic stress but there has been a lack of literature comparing the best osmoprotectant for future breeding programs. In this study, three concentrations (10mM, 20mM, and 30mM) of trehalose and sucrose were compared to determine which osmoprotectant and concentration best improves mung bean growth under 75mM NaCl stress. Four mung bean seeds were pretreated in each trehalose and sucrose solution for 60 hours and grown hydroponically under NaCl stress for one week. Shoot height, root depth, number of root nodes, number of leaves, and leaf appearance of the seedlings were collected to observe growth. One-way ANOVA for shoot height ($P = 0.471$), root depth ($P = 0.162$), number of root nodes ($P = 0.160$), and number of leaves ($P = 0.253$) showed trehalose and sucrose pretreatment did not significantly improve mung bean sprouting under NaCl stress. These results are limited by the small sample size and time frame which did not accurately portray the effects of NaCl stress and osmoprotectant pretreatment. Future trials in this study will increase sample size and time to collect more reliable data on how trehalose and sucrose can be leveraged to create more salt-tolerant crops for global food security.

INTRODUCTION

Soil salinization affects one-fifth of the world's agricultural land (Zaman 2021). Africa and Asia are the most affected, with 122.9 and 193.8 million hectares of agricultural land damaged by soil salinity respectively (Butcher 2016). For most of Earth's history, salts have accumulated in soils through the gradual weathering of rocks, a process known as primary salinization (Mazhar 2022). However, human activity over the past centuries has significantly increased salt accumulation in soil. Secondary salinization caused by anthropogenic activity is primarily a result of three phenomena: sea level rise, drought, and brackish irrigation water.

Ninety percent of heat trapped by anthropogenic greenhouse gases is absorbed by the ocean (NASA 2020). Water expands at higher temperatures which leads to sea level rise (SLR). The warmer water combined with increasing atmospheric temperatures induces the melting of ice sheets and glaciers which further contributes to the increasing volume of seawater. Consequently, the rate of global SLR has doubled over the past one hundred years. By 2100, global mean sea level will rise by 0.5 to 1 meter (Arias 2021) which will submerge thousands of hectares of agricultural land. Besides the direct loss of land, SLR also alters soil composition, creates waterlogging, salinizes

groundwater, and increases storm intensity (Mazhar 2022). The salt deposition from SLR displaces essential nutrients in soil and decreases soil water potential. As a result, crops grown in saline soils experience decreased water intake, ion toxicity, and nutrient deficiency.

Another pressing problem of anthropogenic climate change is drought. The frequency and duration of droughts has increased by 29% since the year 2000 (United Nations 2022). Drought is a major threat to agriculture because the reduced freshwater availability prevents the adequate leaching of saline soils (Hasanuzzaman 2012). Drought is accompanied by high temperatures which increases the rate of water evaporation from soil. Over time, the evaporated water leaves behind increasing concentrations of salt in the soil. Besides drought, groundwater depletion, salinization or contamination also causes freshwater shortages. Freshwater shortages force farmers to use brackish water for irrigation which further exacerbates the problem of soil salinization (Helalia 2021).

Salt stress in saline soils is divided into two stages: immediate hyperosmotic stress and chronic hyperionic stress (Balliu 2016). In salt-induced hyperosmotic stress, plants experience reduced water uptake in their roots which leads to cell membrane disruption, reduced reactive oxygen species (ROS) scavenging capacity, decreased levels of photosynthesis, and decreased stomatal aperture (Gupta 2014). Conversely, in hyper ionic stress, plants accumulate toxic levels of Na^+ and Cl^- ions over time. Na^+ ions are known to compete with K^+ ions for entry into root cells because they are transported with the same ion channels (Keisham 2018). K^+ deficiency leads to disruptions in

photosynthesis, osmoregulation, enzyme activation, protein synthesis, and ion homeostasis (Hafsi 2014). High levels of Cl^- ions significantly reduce plant growth and water use efficiency (Hasanuzzaman 2012). Both Na^+ and Cl^- induce the production of ROSs such as superoxide anions.

One of the most promising coping mechanisms for plants under salt stress is the accumulation of osmoprotectants. Osmoprotectants, also known as compatible solutes, are small, neutrally charged compounds which do not interfere with cellular metabolism even at high concentrations (Jimenez-Arias 2021). Osmoprotectants have the ability to enhance plant metabolism and reduce oxidative damage under abiotic stress.

Carbohydrate sugars are one of the main classes of osmoprotectants. Sugars such as sucrose, glucose, fructans, trehalose, and mannitol have been identified as effective osmoprotectants (Jimenez-Arias). Trehalose is the most extensively studied sugar osmoprotectant. Exogenously applied trehalose has been shown to improve plant growth by promoting K^+ influx and Na^+ efflux, increasing ROS scavenging capacity, decreasing the rate of formation of ROS, increasing the levels of other osmoprotectants such as proline, and improving the activity of stress-related enzymes (Yuan 2022). However, the proper dosage and application method of trehalose needs to be further investigated. Genetically engineered plants with upregulated expression of trehalose experienced negative effects of trehalose toxicity (Hasanuzzaman 2012). Additionally, some studies indicate pretreating seeds with trehalose is more effective than exogenous application of trehalose in irrigation water (Zeid

2009). Besides trehalose, the osmoprotective qualities of other carbohydrate sugars should also be characterized. Sucrose is one of the sugars that has been studied less frequently. In bacteria, sucrose is known to recruit endogenous osmolytes such as glutamate and N-acetylglutaminyglutamine amide during salt stress (Gouffi 1998). In plants, sucrose has been found to reduce intracellular water potential to help restore osmotic equilibrium in plants under salt stress. Several studies on *Arabidopsis thaliana* and triticale have shown that exogenous sucrose can improve crop growth under salt stress (Qiu 2014, Wang 2019). However, those studies only tested one concentration of sucrose. Similar to trehalose, sucrose can also be detrimental in high concentrations because high levels of sugar can inhibit photosynthesis (Wang 2019). Understanding the best dosage and application methods of osmoprotectants such as trehalose and sucrose would help create salt-tolerant crops to combat the rising threat of soil salinization to global food security.

Mung bean (*Vigna radiata*) is an important legume grown primarily on small farms in South, East, and Southeast Asia. It is rich in carbohydrates, iron, and protein which supports the nutritional needs of many Asian communities (Ong 2023). Given that mung bean is a low-altitude crop (Akpapunam 1996), its future production is threatened by sea level rise and soil salinization. There has been a lack in osmoprotectant studies performed on mung beans. Hence, investigating mechanisms of salt stress response for mung beans would help maintain food security for various Asian countries.

MATERIALS AND METHODS

Seed Pretreatment

Four ungerminated mung beans were assigned to each of the following treatment groups: negative control, positive control, 10mM trehalose, 20mM trehalose, 30mM trehalose, 10mM sucrose, 20mM sucrose, and 30mM sucrose. The mung beans in the trehalose or sucrose treatments were pretreated for 2 days with 20mL of their respective solutions (Figure 1). Control group mung beans were pretreated for 2 days with 20mL of distilled water.

Seedling Growth

Each pretreated mung bean was planted in a hydroponic system consisting of a 500mL plastic water bottle and a 1-inch plastic net pot filled with coconut coir (Figure 2). The plastic water bottles were sterilized with 3% hydrogen peroxide and spray painted black to prevent microbial and algal growth. Each bottle was filled with 460mL of nutrient solution containing distilled water and Jack's Classic All Purpose 20-20-20 Water Soluble Plant Food with Micronutrients. The nutrient solution contained 0.5g of fertilizer per 2L of solution. 75mM sodium chloride was added to the nutrient solutions of the positive control, 10mM trehalose, 20mM trehalose, 30mM trehalose, 10mM sucrose, 20mM sucrose, and 30mM sucrose groups. All planted mung beans were grown for 1 week under a Spider Farmer SF600 LED grow light with a 16-hour photoperiod.

Data Collection

After 1 week of growth, each mung bean was removed from its net pot and coconut coir was washed off from the seedling with tap water. A ruler was used to measure shoot height and root depth in

centimeters. The number of root nodes and number of leaves were counted manually for each mung bean.

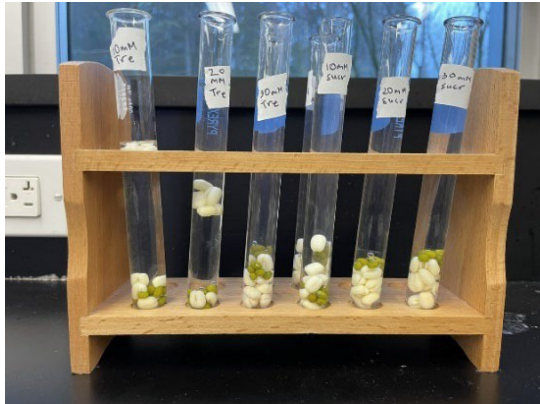


Figure 1: Mung bean seeds germinating in pretreatment solutions.



Figure 2: Mung bean seedlings growing in hydroponic system.

RESULTS

No treatment group presented a significant improvement in growth regarding shoot height, root depth, number of root nodes, and number of leaves. One-way analysis of variance performed for shoot height ($P = 0.471$), root depth ($P = 0.162$), number of root nodes ($P = 0.160$), and number of leaves ($P = 0.253$) all yielded insignificant P -values at $\alpha = 0.05$. The raw data, averages, and photos of final growth are displayed in Table 1, Table 2, Table 3, Table 4, and Figure 3 below.

Table 1: Shoot Height Measurements in Centimeters for Each Treatment

Sample	NC	PC	10mM Tre	10mM Sucr	20mM Tre	20mM Sucr	30mM Tre	30mM Sucr
1	N/A	7.1	8.1	3.2	7.2	2.1	5.1	6.2
2	N/A	3.8	2.6	8.3	5.1	8.8	N/A	3.3
3	N/A	N/A	N/A	N/A	N/A	7.6	N/A	1.2
4	N/A	N/A	N/A	N/A	N/A	3.3	N/A	N/A
Average	N/A	5.5	5.4	5.8	6.2	5.5	5.1	3.6

NC – Negative Control; PC – Positive Control; Tre – Trehalose; Sucr – Sucrose; N/A – Mung beans that did not grow after pretreatment; Average – Not rounded for significant figures, Samples marked N/A not counted in calculation of average.

Table 2: Root Depth Measurements in Centimeters for Each Treatment

Sample	NC	PC	10mM Tre	10mM Sucr	20mM Tre	20mM Sucr	30mM Tre	30mM Sucr
1	0.5	1.4	1.1	1.2	1.4	1.9	1.9	2.5
2	0.3	2.6	1.1	3.5	2.2	2.9	N/A	4.4
3	1.4	N/A	N/A	N/A	N/A	3.5	N/A	2.2
4	N/A	N/A	N/A	N/A	N/A	1.4	N/A	N/A
Average	0.7	2.0	1.1	2.4	1.8	2.4	1.9	3.0

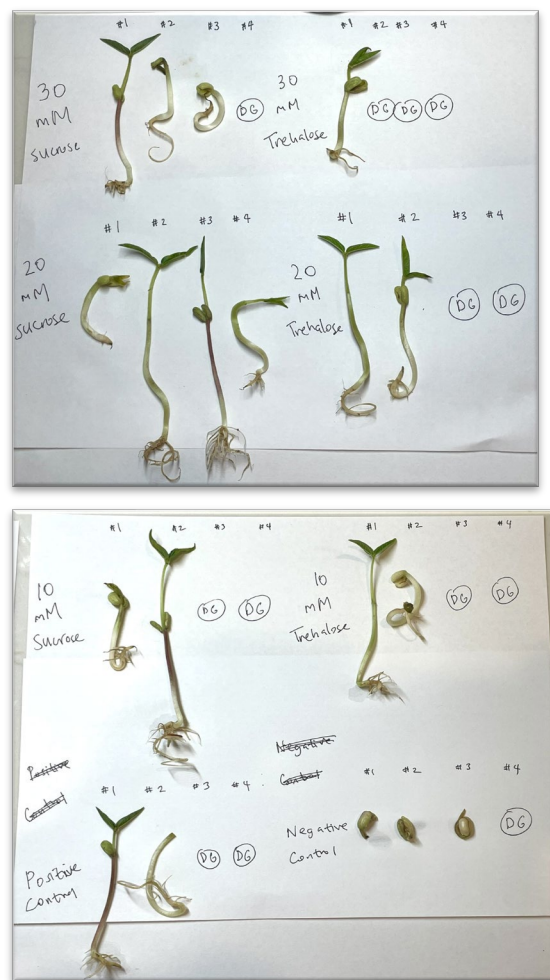
Table 3: Root Node Count for Each Treatment

Sample	NC	PC	10mM Tre	10mM Sucr	20mM Tre	20mM Sucr	30mM Tre	30mM Sucr
1	N/A	2	2	2	2	2	2	2
2	N/A	0	0	2	2	2	N/A	1
3	N/A	N/A	N/A	N/A	N/A	2	N/A	0
4	N/A	N/A	N/A	N/A	N/A	2	N/A	N/A
Average	N/A	1	1	2	2	2	2	1.5

Table 4: Leaf Count for Each Treatment

Sample	NC	PC	10mM Tre	10mM Sucr	20mM Tre	20mM Sucr	30mM Tre	30mM Sucr
1	1	20	11	12	10	7	4	14
2	1	10	6	27	15	18	N/A	3
3	1	N/A	N/A	N/A	N/A	17	N/A	7
4	N/A	N/A	N/A	N/A	N/A	11	N/A	N/A
Average	1	15	8.5	19.5	12.5	13.25	4	8

Figure 3: Mung Beans After 1 Week of Growth



DG – Didn't grow after pretreatment.

ACKNOWLEDGEMENTS

I would like to thank my mentor Dr. Jonathan Lynch at Pennsylvania State University in reviewing my research plan and offering advice to refine my procedure.

DISCUSSION

There is no significant difference in shoot height, root depth, number of root nodes, and number of leaves for salt-stressed mung bean seedlings pretreated with trehalose or sucrose. However, the results of this study are limited by the small sample size and short growth time. Qualitative observations of leaf appearance indicated the mung bean seedlings were not salt stressed. Future studies with larger sample sizes and longer growth times will be needed to collect more accurate data on the osmoprotective effects of trehalose and sucrose on mung beans under salt stress.

If future studies yield significant results, the comparison between the osmoprotective effects of trehalose and sucrose can provide insight on which carbohydrate osmoprotectant is worth the most financial investment for developing salt tolerant crops through commercial seed pretreatment programs, traditional breeding, or genetic engineering.

REFERENCES

- Zaman, Mohammad & M. Exner. (2021). Combatting Soil Salinisation Using Nuclear Techniques: The IAEA Commemorates 2021 World Soil Day. *International Atomic Energy Agency*.
<https://www.iaea.org/newscenter/news/combattling-soil-salinisation-using-nuclear-techniques-the-iaea-commemorates-2021-world-soil-day#:~:text=Globally%2C%20932%20million%20hectares%20of%20agricultural%20land%20%E2%80%93,which%2062%20million%20hectares%20of%20soil%20are%20irrigated>. Accessed 17 Jan 2024.
- Butcher, Kirsten et al. (2016). Soil Salinity: A Threat to Global Food Security. *Agronomy Journal*, 108(6), 2189-2200. <https://doi.org/10.2134/agronj2016.06.0368>. Accessed 17 Jan 2024.
- Mazhar, Sadat et. al. (2022). Impacts of salinization caused by sea level rise on the biological processes of coastal soils – A review. *Frontiers in Environmental Science*, 10. <https://doi.org/10.3389/fenvs.2022.909415>. Accessed 17 Jan 2024.
- NASA & JPL-Caltech. (2020). Rising Tides; Understanding Sea Level Rise. *NASA: Global Climate Change Vital Signs of the Planet*. https://climate.nasa.gov/climate_resources/199/rising-tides-understanding-sea-level-rise/. Accessed 17 Jan 2024.
- Arias, Paola A. & et al. (2021). Technical Summary. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 77. <https://dx.doi.org/10.1017/9781009157896.002>. Accessed 29 Dec 2023.
- United Nations. (2022). World ‘at a crossroads’ as droughts increase nearly a third in a generation. *United Nations*.
<https://news.un.org/en/story/2022/05/1118142#:~:text=And%20from%201998%20to%202017,to%20severe%20and%20prolonged%20droughts>. Accessed 17 Jan 2024.
- Hasanuzzaman, Mirza & K. Nahar, M. Fujita. (2012). Plant Response to Salt Stress and Role of Exogenous Protectants to Mitigate Salt-Induced Damages. *Ecophysiology and Responses of Plants under Salt Stress*, pp. 25-87. https://link-springer-com.stanford.idm.oclc.org/chapter/10.1007/978-1-4614-4747-4_2#Sec12. Accessed 17 Jan 2024.
- Helalia, Sarah A. et al. (2021). Impact of Drought and Changing Water Sources on Water Use and Soil Salinity of Almond and Pistachio Orchards: 1. Observations. *Soil Systems*, 5(3), 50. <https://doi.org/10.3390/soilsystems5030050>. Accessed 17 Jan 2024.
- Balliu, A. & G. Sallaku, T. Nasto. (2016). The effects of some osmoprotectant compounds on growth parameters of pea plants (*Pisum sativum*) under saline conditions. *Acta Horticulture*, 1142, 377-382. <https://doi.org/10.17660/ActaHortic.2016.1142.57>. Accessed 17 Jan 2024.
- Gupta, Bhaskar & B. Huang. (2014). Mechanism of Salinity Tolerance in Plants: Physiological, Biochemical, and Molecular Characterization. *International Journal of Genomics*.
<https://www.hindawi.com/journals/ijg/2014/701596/>. Accessed 17 Jan 2024.
- Keisham, Monika & S. Mukherjee, S. C. Bhatla. (2018). Mechanisms of Sodium Transport in Plants—Progresses and Challenges. *International Journal of Molecular Sciences*, 19(3), 647. <https://doi.org/10.3390/ijms19030647>. Accessed 27 Dec 2023.
- Hafsi, Chokri & A. Debez, C. Abdelly. (2014). Potassium deficiency in plants: effects and signaling cascades. *Acta Physiologiae Plantarum*. <https://doi.org/10.1007/s11738-014-1491-2>. Accessed 17 Jan 2024.
- Jimenez-Arias, David et al. (2021). A Beginner’s Guide to Osmoprotection by Biostimulants. *Plants*, 10(2), 363. <https://doi.org/10.3390/plants10020363>. Accessed 17 Jan 2024.

- Yuan, Gaopeng et al. (2022). Transcriptomic and Metabolomic Analysis of the Effects of Exogenous Trehalose on Salt Tolerance in Watermelon (*Citrullus lanatus*). *Cells*, 11(15), 2338. <https://doi.org/10.3390/cells11152338>. Accessed 17 Jan 2024.
- Zeid, I. M. (2009). Trehalose as Osmoprotectant for Maize Under Salinity-Induced Stress. *Research Journal of Agriculture and Biological Sciences*, 5(5), 613-622. <http://www.aensiweb.net/AENSIWEB/rjabs/rjabs/2009/613-622.pdf>. Accessed 17 Jan 2024.
- Gouffi, Kamila et al. (1998). Sucrose Is a Nonaccumulated Osmoprotectant in *Sinorhizobium meliloti*. *Journal of Bacteriology*, 180(19), 5044-5051. <https://doi.org/10.1128%2Fjb.180.19.5044-5051.1998>. Accessed 17 Jan 2024.
- Qiu, Z. B. et al. (2014). Exogenous sucrose can enhance tolerance of *Arabidopsis thaliana* seedlings to salt stress. *Biologia Plantarum*, 58, 611-617. <https://link.springer.com/article/10.1007/s10535-014-0444-3>. Accessed 17 Jan 2024.
- Wang, L. H. et al. (2019). Effects of exogenous glucose and sucrose on photosynthesis in triticales seedlings under salt stress. *Photosynthetica*, 57(1), 286-294. <https://doi.org/10.32615/ps.2019.030>. Accessed 17 Jan 2024.
- Ong, Pei-Wen & et al. (2023). Environment as a limiting factor of the historical global spread of mungbean. *eLife*, 12. <https://doi.org/10.7554/eLife.85725>. Accessed 15 Jan 2024.
- Akpanunam, M. (1996). Mung bean (*Vigna radiata* (L.) Wilczek). *Food and Feed from Legumes and Oilseeds*. Springer, Boston, MA, USA, pp. 209-215. https://link-springer-com.stanford.idm.oclc.org/chapter/10.1007/978-1-4613-0433-3_23#citeas. Accessed 15 Jan 2024.

The Antibacterial Effect of Ohi'a Lehua Blossom Honey on *Escherichia coli* and *Bacillus cereus*

Eda N. Keskin

Experiment conducted in the Natural Science Department Microbiology Lab at Lake Washington

Institute of Technology under the supervision of Ph.D. E. Barçın Acar

ABSTRACT

Ohi'a Lehua blossoms (OLB), which are critical to honey production and habitats for endangered birds in Hawaii, are currently in danger of extinction due to fungal infection. Honey has been traditionally used as medicine, and providing insight on the potential of OLB honey could help increase awareness for protecting the trees in the future. This investigation evaluates the antimicrobial potential of OLB honey in comparison to Manuka honey (positive control). First, agar well diffusion method was used in triplicate to measure the clear zone of inhibition (ZOI), and evaluate Minimum Inhibitory Concentrations (MIC) with 100, 75, 50 and 25% concentrations of honeys and agave syrup (positive control). Manuka honey showed better antibacterial activity than OLB honey with about an average of 2 mm bigger ZOI for *B. cereus* in 100% concentration. OLB honey showed no ZOI for *E. coli*, while Manuka honey showed an average of 9.3 mm ZOI for 100% concentration (p value = 0.001). Manuka honey showed an MIC of 75% concentration for both bacteria, and OLB honey showed an MIC of 100% for *B. cereus*. Next, broth dilution method using dilutions 1:1, 1:2, 1:4, and 1:8 was done with Manuka and OLB honeys to assess MIC and Minimum Bactericidal Concentrations (MBC). MIC for Manuka honey for both types of bacteria was seen in 1:4 dilution. OLB honey showed an MIC of 1:2 for both bacteria, with more inhibition for *E. coli*. There were no bactericidal effects. Both honeys can be considered as a good candidate for alternate antimicrobial agents in medicine, however OLB honey had less antibacterial ability than Manuka honey. Phytochemical properties of both honeys should be further researched, along with antifungal and antioxidant properties.

INTRODUCTION

This investigation seeks to find whether Ohi'a Lehua blossom (OLB) honey has antibiotic properties caused by its unique compounds, and identify its bactericidal properties. Compounds with bactericidal properties effectively kill bacteria by interfering with or preventing cell wall synthesis, and has the potential to aid in treating wounds and prevent infections. Bactericidal shouldn't be mixed

up with bacteriostatic, which stops the growth of the bacteria rather than kill it.

Although antimicrobial agents decrease the threats posed by microbes, the increasing antibiotic resistance in pathogens causes those agents to be less effective as the pathogens gain tolerance. The excessive use of antibiotics and the resulting antibiotic tolerance in the world makes it difficult to eradicate infectious diseases, especially in underdeveloped and developing countries (Patel and

Chauhan, 2017). However, the danger of antibiotic resistant bacteria isn't exclusive to those countries. Every year, antibiotic resistant bacteria and fungi cause more than 2.8 million infections and 35,000 deaths every year in the United States (CDC, 2019). Therefore, research on novel antimicrobial agents, such as honey, has increased in importance over the years.

Honey has been used to treat wounds and burns since ancient times due to its therapeutic ability to inhibit bacterial growth with its high sugar content and the synergistic effect of its many components (Mandal et al., 2011). It can be applied directly on the wound to prevent inflammation or eaten to boost the immune system. In addition to honey's antibacterial potency, bacteria struggle to develop resistance to honey, making it more significant as a medicine (Bischofberger et al., 2021).

Researchers used various methods to evaluate the antibacterial potential of different types of honey, which include agar disk diffusion, agar well diffusion, and broth dilution methods. Zone of inhibition from Kirby-Bauer testing (agar disk diffusion method) measures how well the antibiotic covered disk prevented bacterial growth around the disk on the agar media. The ZOI is a clear diameter around the disk, and bigger ZOI signifies more antibacterial activity. Agar well diffusion works similarly; a well of six mm in diameter is punched out and then filled in with desired compound or antibiotic. The diameter of the ZOI around the outer surface of the well is then measured. This method is often used to evaluate the antimicrobial activity of plants or microbial extracts. Both agar disk and well

diffusion methods rely on the diffusion of the compound in the agar media. For broth dilution method, bacteria are inoculated into a liquid growth medium in the presence of different concentrations of an antimicrobial agent. Growth is assessed after incubation for a defined period of time, and the Minimum Inhibitory Concentration (MIC) value is read. Decreased turbidity signifies less bacteria growth, and more effective antibacterial effect. Lower MIC means that a lower concentration of the compound of interest is needed in order to show antibacterial effect, meaning that the compound is more effective.

Manuka honey has been researched extensively and was found to have antibacterial properties on both Gram negative and positive bacteria (Lusby et al., 2005). Manuka honey has a dark amber color and is derived from the *Leptospermum* trees of New Zealand. For this famous honey, Wedderspoon company developed a KFactor to measure the pollen content in the Manuka honey. KFactor 16 means the honey is made completely from Manuka flowers, meaning that the honey is monofloral. Monofloral Manuka honey showed to have the lowest MIC when tested against *Staphylococcus aureus*, a Gram positive bacterium, in comparison to the multifloral Manuka/Kanuka blend and clover honey (Lu, 2014). KFactor 16 Manuka honey is therefore more effective in comparison to KFactor 12 Manuka honey.

Honey has many reasons for its antibacterial properties: its low pH value, high sugar concentration, phenolic content, and other compounds present in the honeys. An example of compounds seen in Manuka honey include hydrogen

peroxide, dihydroxyacetone, and methylglyoxal (MGO). MGO binds to the proteins on the flagella and fimbriae and therefore causing loss of structural integrity and subsequently function (Rabie, 2016). This can contribute to the bacteria's death. Methylglyoxal and dihydroxyacetone are mostly seen in Manuka honey, making its properties unique, and antibacterial potential more effective.

Manuka honey isn't the only type of honey with antibacterial effects, however. For example, extra white colored honey called Majra in the Asir region of Saudi Arabia (with 80% concentration) showed a zone of inhibition (ZOI) of 8.67 (± 0.6) mm and 10.67 (± 1.5) mm against *E. coli* and *S. epidermidis*, respectively (Gramh et al., 2018). Amber colored honey had a bigger ZOI than white and extra white honey on both Gram negative and positive bacteria. One of those amber colored honeys (Dharm honey) resulted in having the lowest MIC compared to other colored honeys (Gramh et al., 2018). However, the antibacterial potency of extra white colored honey shouldn't be overlooked as it still has significant antibacterial potency.

One type of honey that is yet to be researched in terms of its antibacterial and chemical components is Ohi'a Lehua blossom honey. Ohi'a Lehua (*Metrosideros polymorpha*) grows in the six islands of Hawaii and can grow in difficult (or almost no) soils, like former lava flows (Staake, 2020). This beloved tree and its blossoms are essential for the survival of many endangered birds species, such as the 'Akepa (Richards). Unfortunately, more than a million of the Ohi'a Lehua blossom's species have died from one of the two fungal species: *Ceratocystis huliohia* or

Ceratocystis lukuohia (Allen, 2021). This spreading disease is called the Rapid Ohi'a Death, which turns all the leaves brown within a couple days and kills the tree in a couple of weeks. These fatal species brought by the tourists of the island can be stopped with increased awareness and precautions, such as cleaning their shoes in hot water and bleach bath before and after entering forests.

Data on the antibacterial properties of the viscous, water white colored Ohi'a Lehua blossom honey are limited, therefore making it important to research the potential. In addition to expanding research on honey, research on OLB honey could serve as a way to increase awareness of the blossom's current situation and encourage tourists to be more careful. If this research finds an exceptional antibacterial property, it'd be easier to advocate for the blossom that grows on recently cooled lava.

This study hypothesizes that when agar well diffusion antimicrobial susceptibility testing is done with Ohi'a Lehua blossom honey, the average clear zone of inhibition against *E. coli* and *B. cereus* bacterial lawns will be at least 8 mm for 75% concentration, showing antibacterial effects. Additionally, the minimum concentration needed to have antibacterial effect for OLB honey will be 100% for both bacteria. When broth dilution method is used, MIC will be achieved at a minimum of 1:1 dilution of OLB honey and Luria Broth media for both bacteria.

MATERIALS AND METHODS

2.1. Honey samples and syrup

Ohi'a Lehua blossom honey is purchased from Big Island Bee company in Kealahou. The

honey is sourced in Hawaii. Raw Monofloral Manuka honey will be used as a positive control and is purchased from Wedderspoon. Honey is sourced in New Zealand, and it has a KFactor of 16. Both Manuka and Ohi'a Lehua blossom honey are stored at room temperature until use. Honey was transported to the lab in glass bowls with a tight plastic lid. Wholesome blue agave syrup by Green Bottle was used as sugar control since it has the same sugar concentration as the honeys (16g sugar per 21g). It was stored in its plastic container and was stored at room temperature.

2.2. Bacterial isolates

Bacterial strains used in this study were obtained from Lake Washington Institute of Technology, WA, U.S.A. Bacteria used were *Escherichia coli* Lac + ATCC 25922 (strain Seattle) and *Bacillus cereus* pellet ATCC 14579. Bacteria were stored at 4°C.

2.3. Agar media preparation

Luria agar media was prepared according to the instructions of the manufacturer with some modifications. It was prepared with 7.64 grams of Difco Luria Agar Base and 250 mL of distilled water. After mixing thoroughly using a magnetic stirrer, solution was put in the microwave until it boiled. It was then autoclaved at 121°C for 15 minutes. While pouring the solution in the petri plates, it was poured slightly thicker than only covering the surface of the plate (resulting in well depths of 4 to 6 mm) in order to store the media longer (Fig. 1 and Fig. 2). The lids were covered right after pouring to prevent contamination due to the flow of air in the cabinet. Agar plates were stored at 4°C.



Figure 1: Petri plates filled with LB agar media cooling in the biosafety cabinet after the media was poured slightly more than just covering the surface.

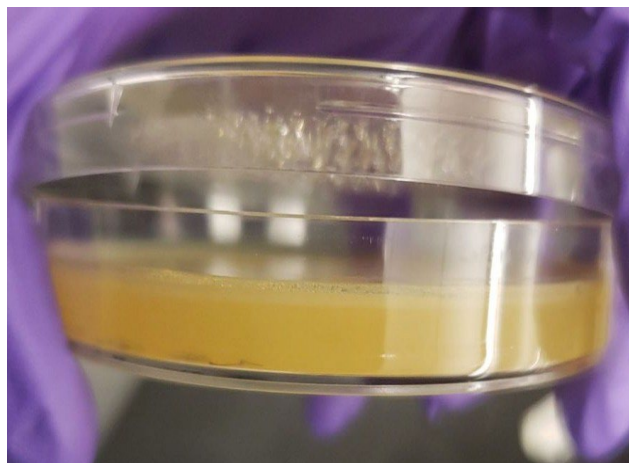


Figure 2: Side profile of a petri plate used for agar well diffusion testing with *E. coli* and 25% compound concentrations (replicate #2). Lighter color seen at top is the grown bacterial lawn, and darker color is the Luria agar media of 5 mm depth.

2.4. Agar disk-diffusion assay for antimicrobial susceptibility testing

Antibacterial activity of honey samples was determined using disk-diffusion test with some modifications. Before spreading bacterial lawns, broth with bacteria were well mixed using a vortex mixer. A total of 24 bacterial lawns on petri plates of *E. coli*, and *B. cereus* were spread using sterile metal loop - 12 petri plates for each bacterium. There were a total of 8 petri plates for each honey compounds tested (Manuka honey, OLB honey), 4 petri plates for negative control and 4 petri plates for agave syrup control. Two replicates of each combination of compound and bacteria were done. The petri plates were rotated 60° two times using a lazy Susan to

spread the bacterial lawn (Fig. 3.). Filter paper disk by Zoro was dipped in honey, and excess honey was allowed to drip off. Disks were placed on the media in a button formation (Fig. 4) for *E. coli* and was repeated with the petri plate for *B. cereus*. Between each time a new disk was dipped and placed, forceps used were sterilized with 70% alcohol and the use of a bacti-cinerator, and the alcohol was washed off. The same two steps were done with the other honey, and agave syrup. Negative controls were paper disks with no compound placed on media with bacteria.

All petri plates were incubated at 30°C for 24 hours. Longest ZOI diameter (in mm) for each disk was measured using a ruler and then recorded.

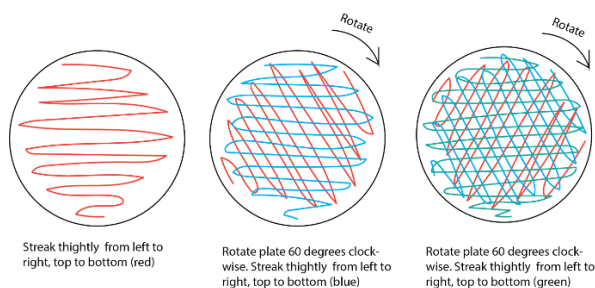


Figure 3: Steps to streaking a bacterial lawn (left to right).

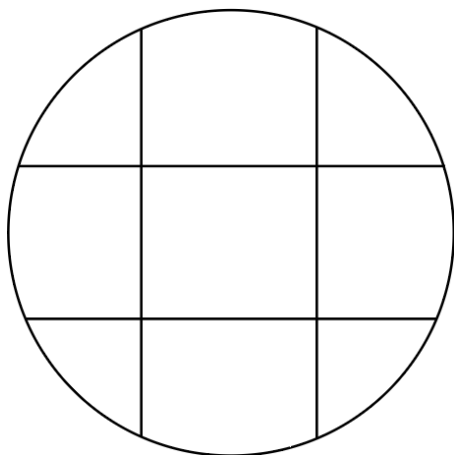


Figure 4: Top 3x3 grid with equal spacing on the intersection points, and are away from the edges of the lid. The intersection points are where the honey disks are meant to be placed

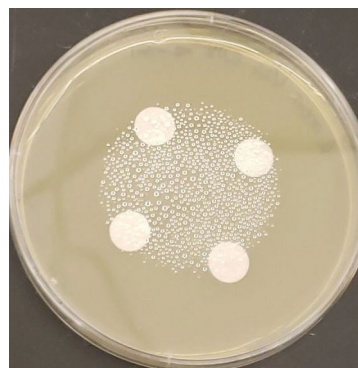


Figure 4B Bottom Media with four filter paper disks placed.

2.5. Agar well diffusion for antibacterial activity testing

Solutions of 100, 75, 50, and 25% (v/v) concentrations were prepared for agave syrup, Manuka honey, and OLB honey using sterile distilled H₂O in 2 mL beakers. The non-pasteurized honeys were warmed up to 36°C in their container to liquify them. Distilled H₂O was filled first, and compound (Agave Syrup, Manuka honey, or OLB honey) filled the container to the maximum. Each container was labeled with concentration and A, M, or L for Agave Syrup, Manuka honey, and OLB honey respectively, and the appropriate % concentration. Before use, solutions were mixed with wooden stick and vortex mixer until there were no visible honey particles remained.

Agar well diffusion method was performed after labeling each petri plate with bacteria and % concentration. First, bacterial lawn for each bacterium was streaked on media. 6 mm diameter wells were punched out on the agar media using the back of pipette tip. Wells used were 4 to 6 mm deep,

depending on the depth of the agar media. Each plate had 3 wells for each compound (Manuka honey, OLB honey, and Agave syrup solutions) in a triangular formation (Fig. 5). Wells were filled with 100 to 200 μ L of each solution, depending on how fast the well filled. Wells for the negative control plates were filled with 80 μ L of sterile distilled water. Pipette tip was cut for easier pipetting of viscous solutions (50% and 75% OLB, 75% and 100% Agave and Manuka). For 100% OLB honey, the well was filled as much as possible using the wooden side of swab. After 24-hour incubation in 30° Celsius, clear zone of inhibition was recorded around the wells using a ruler (margin of error: ± 1 mm) (Table 2).

:

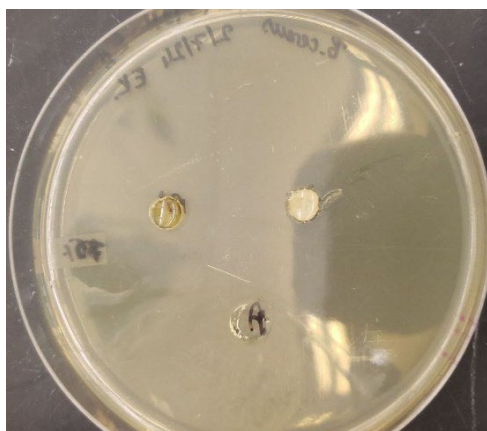


Figure 5: Example petri plate with compounds filled in the wells. Upper left has Manuka honey, upper right has OLB honey, and bottom has Agave syrup of 75% concentration.

2.6. Data processing and statistical analysis for agar well diffusion method

Outliers were either omitted or averaged based on the possible technical errors noted during data collection (Table 2). Replicate 2 of *B. cereus* 50% concentration was omitted due to agar well depth being the thinnest with a depth of 3 mm. Technical errors may include spillage of agave syrup

compound, or errors with pipetting. Results for Manuka and OLB honeys were compared using post-hoc t-test for both bacteria.

2.7. Broth dilution for MIC and MBC value determination

After 0.5 McFarland standard was prepared, bacteria dilutions were determined to match the percent transparency reading from the spectrometer. Diluted honey solutions of 1:1, 1:2, 1:4, and 1:8 (w/v) were prepared for both honey types with Luria Broth in a big batch before putting 5 mL of solution into tubes (Table 1). Unlike for the other solutions, 3 mL of 1:1 dilution solution for both honey types were put into tubes. The tubes were inoculated with 100 μ L of the appropriate bacteria and bacteria dilution per 5 mL of honey and LB solution. Two replicates of the dilution, honey, and bacteria combination were done. Additionally, 1:1 dilutions of both types of honey with no bacteria were prepared as a control. After incubation for 24 hours at 30° Celsius, visible inhibition was determined as the MIC value. Spectrometer readings were taken per each tube. All tubes were plated with sterile disposable swabs: bacterial lawn was spread on Luria agar media to determine bactericidal effects and the number of cultures per plate were counted (Table 3). 300 or more cultures in a plate were recorded to be too many to count (TMTTC). Controls were plated half and half in the same plate.

Table 1: Measurements used for each tube for broth dilution method.

Solution dilution	Honey weight (g)	Luria broth (mL)
1:1	8.72	8.72
1:2	8.34	16.67
1:4	6.25	18.75
1:8	3.125	21.875

RESULTS

3.1. Agar disk diffusion assay

Both Lehua and Manuka honeys showed antibacterial properties with disk-diffusion testing. Results were omitted due to inconsistent values

between replicates and each disk within the same plate.

3.2. Agar well diffusion assay

When agar well diffusion antimicrobial susceptibility testing is done with Ohi'a Lehua blossom honey, the average zone of inhibition against *E. coli* and *B. cereus* bacterial lawns was not at least 8 mm in diameter at 75% concentration, meaning that the hypothesis was rejected. However, both types of honey did show antibacterial properties. For *E. coli* and *B. cereus*, Manuka honey showed a MIC of 75%. OLB honey showed an MIC value of 100% for *B. cereus*.

<i>E. coli</i>												
Concentrations	100%			75%			50%			25%		
Replicates/Compound	1	2	3	1	2	3	1	2	3	1	2	3
Agave	0	0	0	0	0	0	0	0	0	0	0	0
Manuka	9	9	10	0*	10	11	0	0	0	0	0	0
OLB	9*	0	0	0	0	0	0	0	0	0	0	0
<i>B. cereus</i>												
Concentrations	100%			75%			50%			25%		
Replicates/Compound	1	2	3	1	2	3	1	2**	3	1	2	3
Agave	0	0	0	9*	0	0	0	0**	0	0	0	0
Manuka	12	13	10	10	10	8	0	10**	0	0	0	0
OLB	10	10	8	8**	10**	0	9*	11**	0	0	0	0

Table 2: Raw data of the ZOI results (mm) for each type of compound, concentration, and replicate when agar well diffusion method was used.

The results were statistically significant (p -value: 0.001) between Manuka honey and OLB honey for *E. coli* as seen by the post-hoc t -test. During data processing, data was averaged (*) or omitted (**). Italicized data signifies there was inhibition around the well present, but there was a point in the ellipse that the ZOI was zero.

3.3. Broth dilution assay

Visible minimum inhibition was achieved at 1:4 dilution for Manuka honey for both bacteria, and 1:2 for OLB honey for both bacteria, therefore

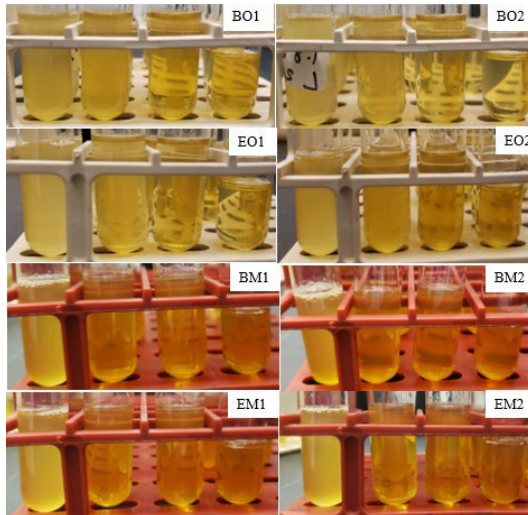


Figure 6: Results of incubation for each tube used for broth dilution method. Photos are in order from OLB honey to Manuka honey, and *B. cereus* to *E. coli*, and from replicate 1 to 2. Dilutions are in order from 1:8 to 1:1 when read from left to right. Photos are labeled with bacteria and honey types (EO: *E. coli* with OLB honey, BO: *B. cereus* with OLB honey, EM: *E. coli* with Manuka honey, BM: *B. cereus* with Manuka honey), and replicate number.

proving the hypothesis correct (Fig. 6). In comparison to the control bacterial lawn, plating results supports conclusions of minimum inhibitory concentrations (Fig. 7).

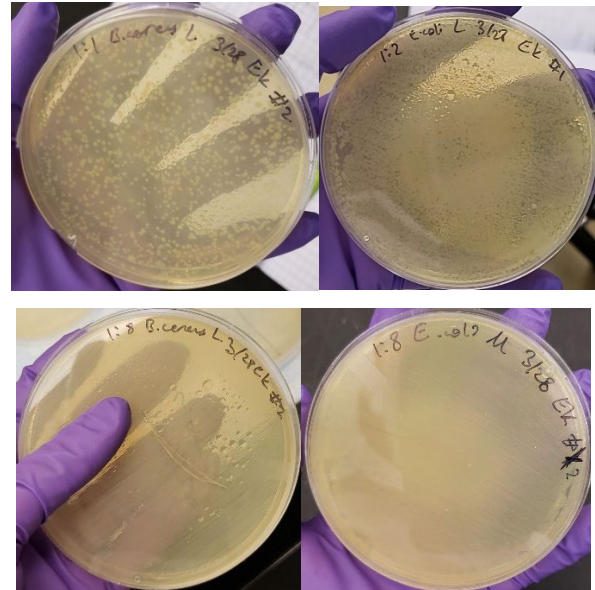


Figure 7: Example plate with inhibition (1st and 2nd from left) in comparison to full bacterial lawn with no inhibition (3rd and 4th from left).

There was no Minimum Bactericidal Concentration seen after plating, as almost all the plates had cultures.

Table 3: Number of cultures on Luria agar media after plating the incubated tubes

Honey type	Manuka honey				OLB honey			
Bacteria type	<i>E. coli</i>		<i>B. cereus</i>		<i>E. coli</i>		<i>B. cereus</i>	
Replicate #	1	2	1	2	1	2	1	2
1:1	3	1	0	2	TMTC	TMTC	TMTC	TMTC
1:2	2	2	2	1	TMTC	TMTC	TMTC	TMTC
1:4	TMTC	TMTC	TMTC	TMTC	TMTC	TMTC	TMTC	TMTC
1:8	TMTC	TMTC	TMTC	TMTC	TMTC	TMTC	TMTC	TMTC

4. DISCUSSION

When agar disk diffusion method was used, the inhibition around the disks of the same compound significantly differed. While some disks had a clear ZOI for OLB honey with *B. cereus*, the others presented no inhibition, or only partial inhibition. This may be due to uncontrollable factors such as the amount of honey that diffused in through the disk, then into the agar media. It may also have resulted from the irregular quantities of honey remaining on the disk once the surplus honey has been dripped off. Additionally, the disks impregnated with Manuka honey moved around the petri plate during incubation, resulting in ambiguous ZOIs. Both OLB and Manuka honeys demonstrated inhibition for both bacteria, so the investigation was taken further with agar well diffusion method.

When agar well diffusion method was used, clear ZOI around the wells were easier to assess. After the data was processed, results showed that for *E. coli* and *B. cereus*, Manuka honey had inhibitory

effects at 75% concentration, while OLB honey showed inhibitory effects for *B. cereus* at 100% concentration. OLB honey showed no ZOI for *E. coli*, while Manuka honey showed an average of 9.3 mm (SD = 0.58) ZOI for 100% concentration. There was a significant difference between OLB honey and Manuka honey results for *E. coli*, meaning that Manuka honey had a significantly bigger antibacterial effect (p value = 0.001). Manuka honey also showed an average ZOI of 10.5 mm (SD = 0.5) for 75% concentration against *E. coli*, which was surprisingly bigger than the average for 100% concentration. This may be due to the slight differences in techniques and human error, since replicate three was conducted in a different week. It could also be because the sample size was too small because of the outlier.

Manuka honey showed a bigger average ZOI for *B. cereus*, with an average of 11.7 mm (SD = 1.53) for 100% concentration, and an average of 9.3 mm (SD = 1.15) for 75% concentration.

Decreasing antibacterial effect as the concentration decreases is due to the decreased concentration of antibacterial compounds in the media, and the decreased osmotic pressure. OLB honey also showed antibacterial effects at 100% concentration with a ZOI of 9.3 mm (SD = 1.15) against *B. cereus*.

A ZOI that is elliptical that at one point touches the borders of the well could signify that the ZOI, despite being clear, may be falsely attributed to the antibacterial properties of the compound in the well. Reasons for the lack of bacterial growth around the well may be due to human error in pipetting and resulting compound spillage, as seen with the agave syrup that seemed to have a ZOI of 9 mm in *B. cereus*. It may also be due to the fact that bacteria didn't properly grow for another reason, such as errors in preparing agar media (e.g. cracks or bubbles in agar).

Along with clear ZOIs, there were also partial inhibition zones seen in petri plates with *E. coli*. These were caused by the osmotic pressure since they were seen with both honey types and agave syrup, and the diameter of the partial zone of inhibition decreased as the concentration of the compound decreased. The osmotic pressure on the bacterial cells caused bacteria to lose water, resulting in the inhibition of the bacteria from growing on the agar media. This data was not considered in evaluating antibacterial effects of honey since it was also seen in the sugar syrup, and the purpose of this investigation is to assess synergistic antibacterial potential of the unique compounds found in the honeys.

The viscosity of the honeys made it difficult to pipette and to prepare solutions, which may've

caused varying results. OLB honey, which crystallizes in room temperature, may've diffused in the agar media less than the Manuka honey, resulting in a smaller average ZOI. In order to reduce confounding variables, the investigation was continued with broth dilution method.

Broth dilution method was more effective for evaluating antibacterial effect due to removing the limitations of differing honey diffusion due to viscosity. Visual MIC's between both replicates of each dilution, honey type, and bacteria type were consistent. Manuka honey showed inhibitory effects at a lower concentration than OLB honey did for both bacteria; Manuka honey needed a minimum dilution of 1:4 while OLB honey needed at least a dilution of 1:2 to show visible bacterial growth inhibition in the tubes. This means that Manuka honey was more effective at inhibiting bacterial growth; however, OLB honey did have antibacterial properties.

When the results were plated, the conclusions about MIC were supported. Inhibition of bacteria on the petri plates was seen at dilutions that matched the determined MIC values for both honey types. There were no complete bactericidal effects (petri plates with no culture growth) when looked at the average of both replicates, meaning that the honeys couldn't completely kill off the bacteria, even though they inhibited its growth. The honeys inhibited bacterial growth more as their concentration increased as seen by the decreasing turbidity and increasing percent transparency.

There was human error involved that resulted in contamination, as seen by 1 culture growing on 1:8 and 1:4 OLB honey controls, and in

1:4 Manuka control. Errors may include unsterile pipetting, or contamination of honeys during transportation.

This research suggests that OLB honey has significant antibacterial properties and should further be researched.

FURTHER IMPLICATIONS

The phytochemical properties of both honeys should be further researched, along with other synergistic antibacterial effects Manuka honey and OLB honey have. In addition, other properties such as antioxidant or antifungal properties could be researched. The honeys should be further researched to develop them into medicine or treatment, such as band aids.

CONCLUSION

Both OLB and Manuka honeys showed antibacterial effect against Gram positive and Gram negative bacteria. Honey was able to prevent the spread and growth of bacteria due to the synergistic

effect of its compounds, molarity, and pH value.

Honey can kill cells or prevent bacterial growth even without its sugar components. Results show that when honeys are used at the appropriate therapeutic level, they can aid with our fight against antibiotic resistant bacteria and provide us alternative options for wound treatment. Additionally, the importance of Ohi'a Lehua trees for the environment and human society has increased, as the honey from the tree is valuable for medical progression.

ABBREVIATIONS

ATTC: American type culture collection

MGO: Methylglyoxal

OLB: Ohi'a Lehua blossom

MIC: Minimum inhibitory concentration

MBC: Minimum bactericidal concentration

ZOI: Zone of inhibition

LB: Luria Broth

TMTC: Too many to count

SD: Standard deviation in millimeters

BIBLIOGRAPHY

“Antibiotic Resistance Threats Report.” Centers for Disease Control and Prevention. (2019).

Centers for Disease Control and Prevention. <https://www.cdc.gov/drugresistance/biggest-threats.html>.

Allen, K. (2021). 5 Things You Need to Know about the ‘ōhi‘a Lehua Tree. Hawaii Magazine.

<https://www.hawaiimagazine.com/5-things-you-need-to-know-about-the-ohia-lehua-flower/>.

Balouiri, M., Sadiki, M., & Ibsouda, S. K. (2016, April). Methods for in vitro evaluating antimicrobial activity: A Review. Journal of pharmaceutical analysis. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5762448/>.

Bischofberger, A. M., Pfrunder Cardozo, K. R., Baumgartner, M., & Hall, A. R. (2021). Evolution of honey resistance in experimental populations of bacteria depends on the type of honey and has no major side effects for antibiotic susceptibility. *Evolutionary Applications*, 14(5), 1314–1327. <https://onlinelibrary.wiley.com/doi/10.1111/eva.13200>.

Gramh, H.A., Khan, K.A., Alshehri, A.M.A. (2018). Antibacterial potential of some Saudi honeys from Asir region against selected pathogenic bacteria. *Saudi Journal of Biological Sciences*, 26(6):1278-1284. <https://doi.org/10.1016/j.sjbs.2018.05.011>. Lu, J., Turnbull, L., Burke, C. M., Liu, M., Carter, D. A., Schlothauer, R. C., Whitchurch, C. B., & Harry, E. J. (2014). Manuka-type honeys can eradicate biofilms produced by *Staphylococcus aureus* strains with different biofilm-forming abilities. *PeerJ Life and Environment*, 2, e326. <https://doi.org/10.7717/peerj.326>.

Lusby, P. E., Coombes, A. L., & Wilkinson, J. M. (2005). Bactericidal activity of different honeys against pathogenic bacteria. *Archives of medical research*, 36(5), 464–467. <https://doi.org/10.1016/j.arcmed.2005.03.038>.

Mandal, M. D., & Mandal, S. (2011). Honey: its medicinal property and antibacterial activity. *Asian Pacific journal of tropical biomedicine*, 1(2), 154–160. [https://doi.org/10.1016/S2221-1691\(11\)60016-6](https://doi.org/10.1016/S2221-1691(11)60016-6).

Patel, A., Chauhan, P.B. (2017). Antimicrobial effect of Honey on MRSA isolated from pus samples. *Int. J. Drug Res. Technol.* 6, 58-63.

Rabie, E. (2016, March 17). *How methylglyoxal kills bacteria: An Ultrastructural study*. *Ultrastructural pathology*. <https://pubmed.ncbi.nlm.nih.gov/26986806/>.

Richards, H. (n.d.). *Endangered species spotlight: ‘ākepa (Loxops Coccineus): U.S. Fish & Wildlife Service*. FWS.gov. <https://www.fws.gov/story/endangered-species-spotlight-akepa-loxops-coccineus#:~:text=Since%20%27ākepa%20nest%20exclusively%20in,and%20their%20rain%20forest%20habitat>.

Staake, J. (2020, April 24). *Ohia Lehua, a legendary plant of Hawaii*. *Birds and Blooms*. <https://www.birdsandblooms.com/gardening/flower-gardening/ohia-lehua-legendary-plant-hawaii/#:~:text=As%20with%20most%20places%20on,directly%20on%20the%20basalt%20rock>.

COMPUTATIONAL BIOLOGY, ARTIFICIAL INTELLIGENCE

A Novel Approach for Detection and Quantification of Stool on Abdominal Radiographs in Children using Machine Learning

Roxana Stanescu

INTRODUCTION

The goal of this project is to create a convolutional neural network (CNN) that can identify and quantify stool on x-rays in order to avoid interobserver variability and aid in diagnosis of constipation in pediatric patients. Over 2000 abdominal radiographs have been performed in 2022 alone at Seattle Children's Hospital to evaluate for stool burden in children with clinical concern for constipation. According to John Hopkins University, around 3% of total clinical visits for preschool and school-age children are related to constipation (Wilsey). Constipation, if left unaddressed, can lead to complications such as fecal impaction, anal fissures, and rectal bleeding. Early intervention through the quantification of stool burden can help prevent these complications.

Radiographs are a powerful tool that can help diagnose constipation, but quantification of stool burden on x-rays is challenging. There have been several attempts to assess stool burden on x-rays, one of the most commonly known being the Leech method that assigns a score from 0 to 15 based on the distribution, size, and density of stool throughout the colon. However, subsequent studies assessing this scoring system found significant inter and interobserver variability and concluded that this method of assessing fecal load has a limited clinical value (Lorjin et. al, 2006). The difficulties arise from interobserver variability in the way stool burden is

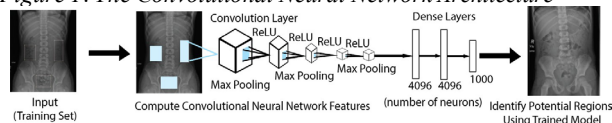
quantified subjectively, where different radiologists evaluating the same abdominal radiograph may quantify the stool burden (the amount of stool in the colon) differently.

Research has shown that machine learning based diagnosis can be a powerful tool to help with this issue. For example, convolutional neural networks, a type of machine learning, have been shown to help diagnose and detect pneumonia on chest radiographs (Kaushik, et al, 2020) and create an imaged based screening tool for COVID-19 on CT scans (Dumakude, Ezugwu, 2023). These advances have made machine learning medical diagnosis an exponentially more useful diagnostic tool. Automated analysis through a machine learning model allows for the generation of quantifiable and standardized results which not only ensures a more objective assessment, but provides a quantifiable, reliable baseline result. The efficient and accurate quantifiable results also contribute to preventing unnecessary additional tests and interventions, further reducing overall healthcare expenses. This would make abdominal radiographs a more efficient, cost-effective, and powerful way to help diagnose constipation in kids. Therefore, the goal of this project is to create a machine learning model that can identify and quantify stool on x-rays in order to avoid interobserver variability and eventually aid in diagnosis of constipation in pediatric patients.

METHODOLOGY

The initial step involved preprocessing the radiograph from the dataset (evaluated and labeled by radiologists at Seattle Children's) including resizing the image dimensions and applying augmentations to make them easier to use. The image dimensions were resized to 500 x 500, changed to a grayscale image with one color channel of gray and made into a dataset in which stool masks and abdominal radiographs are overlaid using Vascal POC to label bounding boxes indicative of stool presence. The dataset was then partitioned into training and testing sets, adhering to an 80:20 ratio with all images.

Figure 1: The Convolutional Neural Network Architecture



Preprocessing of the images begins by uploading the images to the model into a data set. Then, the images are resized to 500 x 500 pixels and transformed into a grayscale picture. Features of the image are then extracted to make it easier for the CNN to detect stool presence. Lastly, images are then used to train the CNN.

The proposed CNN architecture, inspired by prior work on classifying pneumonia on x-rays (Kaushik, et. al, 2020) incorporated convolutional layers, activation functions, pooling layers, dropout layers, and fully connected layers.

First off, in convolutional neural networks, convolutional layers serve as fundamental components for extracting essential features from input images. This process involves the systematic application of convolutional filters to different regions of the image. The horizontal and vertical strides, representing the step size of the filter as it moves across the image, work in feature extraction.

A combination of filter size (known as a kernel), stride, and input dimensions determines the size of the feature map, influencing the granularity of the features extracted. For this model, a kernel of (3,3) was applied because previous studies have shown a smaller kernel size is most accurate for use in medical image reading. Next, an activation function known as Rectified Linear Unit (ReLU), was applied. Activation functions in neural networks introduce non-linearities, enabling the network to learn complex patterns and relationships in data. For this project, ReLU which is used in between convolutions outputs the input directly if positive and zero otherwise.

Following activation, pooling layers were employed to down sample the spatial dimensions of the feature maps produced by convolutional layers, reducing computational complexity and focusing on essential features. Max pooling, the technique used in this project, extracts the maximum value from a set of neighboring pixels in the feature map. These operations retain the most salient information, discarding redundant details. Pooling layers contribute to translation invariance, making the network more robust to variations in spatial orientation. While they reduce dimensionality, they aid in preserving important features for subsequent layers, enhancing the overall efficiency and performance of the neural network.

For the last steps, dropout layers were used as a regularization technique employed during training to prevent overfitting which happens when the model is trained on the data provided but lacks in performance when presented with new data. Dropout layers act as a form of ensemble learning within the

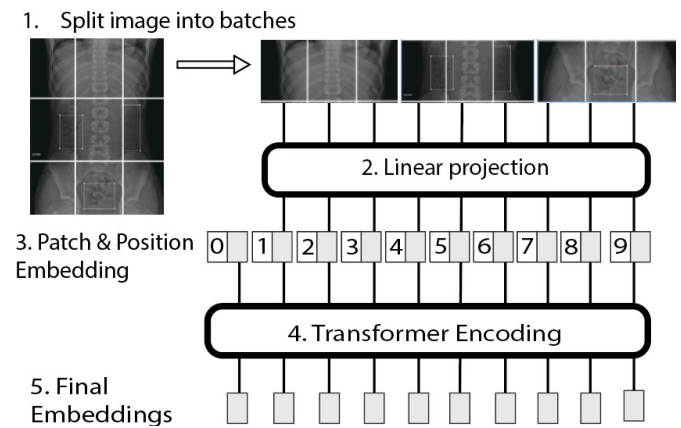
network, improving its robustness, and reducing the likelihood of relying too heavily on specific features. During each training iteration, random neurons are temporarily "dropped out" or ignored, along with their connections, with a probability determined by a dropout rate. This process encourages the network to learn redundant representations, enhancing generalization to unseen data. By introducing variability in the training process, dropout layers contribute to more resilient and adaptable neural networks.

Following these layers, fully connected layers, also known as dense layers, play a critical role in connecting the extracted features from previous layers to the final output layer. Neurons in a fully connected layer are connected to all neurons in the preceding layer, allowing for complex relationships and combinations of features to be learned. This layer is employed as the end of the neural networks for tasks quantification. Lastly, the Sigmoid function will be applied. The Sigmoid function determines the output between 0 and 1, which will be interpreted as the probability of stool presence, and a threshold will be applied to make a binary decision.

Unfortunately, this initial model was unsuccessful. Understanding the appearance of stool on radiographs is multifaceted, as it can be influenced by factors such as its composition, hydration level, and positioning within the gastrointestinal tract. Recognizing stool requires an understanding of how these features interrelate across the entire image. However, convolutional neural networks (CNNs), which excel at capturing local details, may struggle to discern the complex

patterns indicative of stool due to their focus on local rather than global context. Consequently, CNNs may be ineffective in detecting stool. In contrast, Vision Transformer (ViT) offered a promising alternative. By utilizing an attention mechanism, ViT efficiently captures global relationships among features, potentially enabling it to better recognize the intricate patterns associated with stool on radiographs (Al-ammuri, Gebali, Kanan, 2023).

Figure 2: Vision Transformer Model



A Vision Transformer (ViT) is a deep learning model that applies the transformer architecture, originally designed for natural language processing, to image recognition tasks. Instead of relying on convolutional layers like CNNs, ViTs process images by splitting them into fixed-size patches and treating them as sequences of tokens, enabling effective feature extraction and classification.

Unlike Convolutional Neural Networks (CNNs), which rely on convolutional layers, ViTs utilize the transformer architecture originally designed for natural language processing tasks. This approach involves splitting the input image into fixed-size patches, like sequences of tokens, facilitating efficient feature extraction and classification. The process begins with dividing the

image into batches of patches using a sliding window approach, ensuring efficient processing of spatial information. These patches undergo linear projection, transforming raw pixel values into initial embeddings through learned linear projection layers. In this ViT framework, an "embedding" denotes the representation of an image patch after transformer processing, encompassing spatial and contextual information essential for accurate image understanding.

Subsequently, positional embeddings are incorporated to encode spatial relationships among patches within the image grid, facilitating spatial understanding. Following this, the patches, now augmented with positional embeddings, undergo processing through transformer encoder layers. These layers employ self-attention mechanisms to weigh the significance of different patches based on semantic similarities, capturing global dependencies

within the image. Subsequent feedforward neural networks refine attended features to extract higher-level representations, iteratively refining embeddings to capture complex relationships and abstract features. Ultimately, the output of the transformer encoder layers comprises final embeddings representing rich, high-level features of the image, encapsulating both local and global information. These final embeddings, enriched with the model's understanding of the input image, serve as a robust foundation for various downstream tasks, including the detection of stool.

Utilizing COCO, an image labeling software, the images were then labeled with bounding boxes indicating stool presence and were then fed to the Azure Cognitive Services Transformer Model and the model with a training data set of 80 images and validation of 20.

RESULTS

Figure 3: Vision Transformer Model

Trials	Mean Average Precision @ IoU of 30	Mean Average Precision @ IoU of 50	Mean Average Precision @ IoU of 75
Performance for gray-scale images (2-hour training)	93.1%	75.4%	11.5%
Performance for gray scale images (1 hour training)	91.0%	74.7%	23.9%
Performance for color/RGB Images (1 hour training)	77.1%	28.0%	4.2%

Mean Average Precision (mAP) at 93.1% with an IoU of 30 signifies the model is achieving high precision at an Intersection over Union of 30. The model achieves a lower precision on color images due to the fact that RGB images have three color channels opposed to one in the grayscale images, leading to a larger range in pixel values reducing the model's ability to detect stool.

The Mean Average Precision (mAP) metric is a benchmark in computer vision research for evaluating object detection models. The mAP provides an assessment of model efficacy and reliability by quantifying average precision across all classes at a specific Intersection over Union (IoU) threshold, typically denoted as a percentage. For instance, a mAP of 93.1% at an IoU of 30 indicates that the model achieved an average precision of 93.1% when comparing predicted bounding boxes to ground truth boxes, with an IoU threshold of 30%.

IoU measures the extent of overlap between predicted and ground truth regions, serving as a critical factor in assessing detection accuracy. As the confidence threshold increases, restricting predictions to high-confidence labels, IoU will approach 100%, indicating precise object localization. The model's mAP goes down at a higher IoU simply because the model is adept at detecting even what was not labeled so when the model calculates the IoU the additional correctly labeled bounding boxes that were not present in the original image bring the mAP down. The consistent performance on abdominal radiographs beyond the training dataset suggests the model's suitability for real-world deployment.

However, when presented with color/RGB images, the model's performance notably declines. At IoU 30%, the model achieves only 77.1% mAP with 1-hour training, indicating a 16% drop compared to grayscale performance. This significant decrease suggests the heightened complexity introduced by multiple color channels as color images have three (red, green and blue) as compared to grayscale. Since there is a wider range of pixel

values present in any given color image as compared to its grayscale counterpart, this likely had an effect on the training of the model, and it was better able to discern patterns indicative of stool on grayscale images.

Although, in scenarios where the model lacks sufficient training data representation, performance may diminish, highlighting the importance of comprehensive dataset coverage for robust model performance across diverse scenarios. A future step would be to train the data on a larger and more extensively labeled training data set to achieve a higher mAP at an IoU of 75 or greater.

DISCUSSION

The successful development of the detection model has exceeded initial engineering objectives achieving a mean average precision of 93.1% at an IoU of 30, demonstrating the model's efficacy in identifying stool patterns on radiographs. In the future, the model will be retrained on a larger image dataset to see how the model's mAP improves. Building upon the achievement of a successful model, the next phase entails refining the model to translate detection outcomes into an area-based formula, enabling the derivation of a stool burden rating. This enhancement will hopefully provide clinicians with a quantitative measure of stool presence, facilitating more precise clinical assessments. Rigorous testing protocols will be implemented to ensure the reliability and applicability of results before the model's deployment. Following validation, efforts will focus on making the model accessible to radiologists for testing, potentially through integration into a website or another user-friendly platform.

The final goal is to validate this method and obtain necessary approvals for clinical implementation and incorporation of the model into the diagnostic imaging workflows. Continuous refinement and optimization are integral components of this process, guided by ongoing feedback from radiologists and evolving clinical requirements. Moreover, correlating stool burden ratings with patient symptoms and clinical scores of

constipations is a critical aspect of the project's end goal. While quantifying stool presence is valuable for diagnostic purposes, understanding its correlation with patient-reported symptoms is essential for comprehensive patient care. Therefore, the ultimate objective includes establishing robust correlations between stool burden ratings and constipation symptoms, aiming to improve treatment interventions and patient outcomes.

REFERENCES

- Al-hammuri, K., Gebali, F., Kanan, A. et al. Vision transformer architecture and applications in digital health: a tutorial and survey. *Vis. Comput. Ind. Biomed. Art* 6, 14 (2023). <https://doi.org/10.1186/s42492-023-00140-9>
- Azure AI Vision. Azure Vision as a Service | Microsoft Azure. (n.d.). <https://azure.microsoft.com/en-us/products/ai-services/ai-vision>
- de Lorijn, F., van Rijn, R. R., Heijmans, J., Reitsma, J. B., Voskuil, W. P., Henneman, O. D., Taminiau, J. A., & Benninga, M. A. (2006). The Leech method for diagnosing constipation: intra- and interobserver variability and accuracy. *Pediatric radiology*, 36(1), 43–49. <https://doi.org/10.1007/s00247-005-0031-z>
- Dumakude, A., & Ezugwu, A. E. (2023, June 30). *Automated covid-19 detection with Convolutional Neural Networks*. Nature News. <https://www.nature.com/articles/s41598-023-37743-4>
- Hajian-Tilaki K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine*, 4(2), 627–635.
- Joongwon Kim Hankuk Academy of Foreign Studies, Kim, J., Studies, H. A. of F., Hojun Lee Hankuk Academy of Foreign Studies, Lee, H., Taeseon Yoon Hankuk Academy of Foreign Studies, Yoon, T., & Metrics, O. M. A. (2017, November 1). ACM Other conferences. <https://dl.acm.org/doi/10.1145/3168776.3168798>
- Kaushik, S. V., Nayyar, A., Kataria, G., & Jain, R. (April 2020) Pneumonia detection using convolutional neural networks (cnns). https://www.researchgate.net/publication/340961287_Pneumonia_Detection_Using_Convolutional_Neural_Networks_CNNs
- Tlemcen, I. L. U. of, Lahsaini, I., Tlemcen, U. of, Mostafa El Habib Daho University of Tlemcen, Daho, M. E. H., Mohammed Amine Chikh University of Tlemcen, Chikh, M. A., University, L. T., Technology, B. I. of, Sfax, D. research center of, & Metrics, O. M. A. (2020, October 1). *Convolutional neural network for chest X-ray pneumonia detection: Proceedings of the 1st International Conference on Intelligent Systems and Pattern Recognition*. ACM Other conferences. <https://dl.acm.org/doi/10.1145/3432867.3432873>
- Wilsey, M. (2023, July 21). *Dealing with kids and constipation*. Johns Hopkins Medicine. <https://www.hopkinsmedicine.org/health/wellness-and-prevention/dealing-with-kids-and-constipation>

Exploring the Influence of AI-Generated Imagery on E-Commerce Consumer Behavior: A Comparative Study with Traditional Imagery and AI-Generated Imagery

Ahanya John

ABSTRACT

In today's evolving digital sphere, many industries are incorporating AI into daily operations. One industry that has integrated AI is e-commerce, where the way products are presented holds significant influence over consumer decision-making. In 2023, \$25.2 billion was put towards AI startups. With the emergence of AI-generated imagery, many businesses are exploring innovative solutions to engage online shoppers and drive their sales. This study centers on how exposure to AI-generated imagery affects consumer perceptions of product appeal compared to traditional imagery. To address this, participants were randomly assigned to two groups: one exposed to AI-generated imagery and the other to traditional imagery, hypothesizing that AI-generated imagery will show lower levels of visual appeal, perceived trustworthiness, and purchase intent compared to traditional imagery. Each group evaluated product images based on visual appeal, purchase intentions, and perceived trustworthiness using Likert rating scales. The findings are that participants exposed to AI-generated imagery exhibited higher ratings of visual appeal and increased perceived trustworthiness but inconclusive to purchase intentions compared to those who were exposed to traditional imagery. This research provides insights for businesses to optimize marketing strategy and enhance user experience.

INTRODUCTION

This research seeks to understand the nuanced dynamics between AI-generated visuals and traditional presentation methods and its impact on the e-commerce industry. With the sudden rise of artificial intelligence, it has affected many industries of society. From healthcare, to agriculture, and even the e-commerce industry. There are two main types of artificial intelligence, traditional AI, and generative AI. Traditional AI uses preset algorithms to perform tasks, while generative AI can produce different types of content including audio, images, text, and data through prompts given by the user. One big industry in the United States and across the world is e-commerce. In 2022, the e-commerce industry generated 789.01 billion dollars (International Trade Administration, 2022) and E-

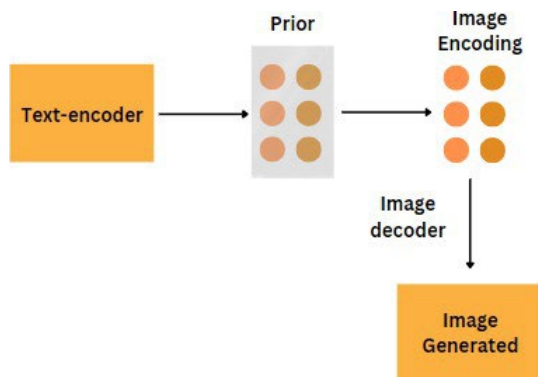
commerce businesses adopting generative AI has grown by 270% in the last four years (Forbes, 2022). Today popular e-commerce sites like Amazon and Alibaba use generative AI as virtual shopping “assistants”. These “assistants” have the ability to analyze past purchases, browse a user’s history to recommend products, and predict user trends (Ihnatchyck, 2023).

In the past 10 years, AI models have been developing. Some notable ones being ChatGPT (developed by OpenAI), Dall-E (OpenAI), and Bard (developed by Google). In 2023, venture capitalists invested \$25.2 billion into generative AI companies and startups.

This research study will use Dall-E (an AI system that can create realistic images and art from a description in natural language) for the experimental

procedure. Dall-E is powered by a diffusion model. Diffusion models work by turning a text prompt into a text encoder which maps the prompt to a presentation space. Then, a model known as “Prior” lays out the text encoding to an encoding of a corresponding image that captures the meaning of the prompt in the encoded text. Finally, an image decoder creates an image based on semantic information.

Figure 1: Dall-E Diffusion Model



Dall-E diffusion model graphic.

The most relevant research that has analyzed the impact of generative AI use on the perceived quality of reviews on popular online platforms, finding that generative AI use drastically reduces the average quality of a reviewer output when incorporated into the review production process (Bart, 2023). Other studies have shown that there is a bias toward AI generated content. One study found that participants in the experiment consistently deprecate art labeled as AI-made relative to human-made art (Iyengar, 2023). But what has not been done is a comparative study with different presentation formats in the realm of e-commerce and product presentation formats. This project seeks to delve into

the dynamics of online shopping behavior by investigation how different presentation formats – AI generated imagery and traditional imagery - affect a consumer's preference, purchase intention, and perceived trustworthiness. In this study it is hypothesized that if online shoppers are exposed to products featuring AI generated imagery, then there will be lower levels of visual appeal leading to decreased ratings of perceived trustworthiness and purchase intentions.

METHODOLOGY

For this experiment, 107 participants were selected based on age (age 14 and above), gender, and frequency of online shopping. To recruit participants, the sampling procedure consisted of sharing recruitment flyers with friends, family, colleagues, and finding participants on Prolific. Before beginning the experiment, participants were asked to fill out an informant consent that established their approval to be a part of the study. Participants data was stored in a secure and password protected file. After all participants were selected, they randomly assigned participants using a random number generator in survey platform Qualtrics, into two different groups. One group to AI-generated imagery, and the second group to the control group with traditional images from Unsplash.com. For the first group (AI- generated), DALL-E was used to generate 7 realistic product images based on user prompts. Participants need to explore images and give ratings on a scale of 1- 7 (worst to best) of visual appeal, purchase intentions, and perceived trustworthiness. Lastly, for control groups, use high-quality traditional product images

from Unsplash.com. Participants explore images and give feedback and rating on a scale of 1-7 of visual appeal, purchase intentions, and perceived trustworthiness.

After all data was collected, data inputs were exported from Qualtrics into IBM SPSS Software

for data analysis. Using the software, descriptive statistics test was used to find mean, median, and mode as well as measures of dispersion (variance, standard deviation, range, minimum and maximum).

Table 1: Survey Questions – **Perceived Trustworthiness**

<i>“Please Indicate to what extent you find Product # _ to be trustworthy”</i> <i>1 – Strongly disagree</i> <i>2- Disagree</i> <i>3 – Slightly Disagree</i> <i>4 – Neutral</i> <i>5 – Slightly Agree</i> <i>6 – Agree</i> <i>7 – Strongly Agree</i>

Table 3: Survey Questions – **Purchase Intentions**

<i>“I would consider purchasing Product #_”</i> <i>1 – Strongly disagree</i> <i>2- Disagree</i> <i>3 – Slightly Disagree</i> <i>4 – Neutral</i> <i>5 – Slightly Agree</i> <i>6 – Agree</i> <i>7 – Strongly Agree</i>
--

Table 2: Survey Questions – **Visual Appeal**

<i>“Please indicate to what extend you find Product # _ to be visually appealing”</i> <i>1 – Very Unappealing</i> <i>2 – Disagree</i> <i>3 – Slightly Disagree</i> <i>4 – Neutral</i> <i>5 – Slightly Agree</i> <i>6 – Agree</i> <i>7 – Strongly Agree</i>
--

Table 4: Traditional Imagery (*Images that were presented to Group 1*)

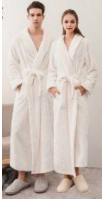
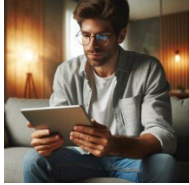
Figure 2.1 — Traditional Shoes	
Figure 2.2 — Traditional Robe	
Figure 2.3 — Traditional Tablet	
Figure 2.4 — Traditional Backpack	
Figure 2.5 — Traditional Plates	
Figure 2.6 — Traditional Chair	
Figure 2.7 — Traditional Lamp	

Table 5: AI Generated Imagery (*Images that were presented to Group 2*)

Figure 3.1 – AI Shoes	
Figure 3.2 – AI Robe	
Figure 3.3 – AI Tablet	
Figure 3.4 – AI Backpack	
Figure 3.5 – AI Plates	
Figure 3.6 – AI Chair	
Figure 3.7 – AI Lamp	

Experimental Conditions Data Tables

Table 6.1: **Visual Appeal** of products; average mean, standard deviation, and p-value

Product	Mean	Standard Deviation	P-value
Shoe	Traditional – 5.0000 AI – 5.0370	Traditional – 1.49358 AI – 1.93115	0.937
Backpack	Traditional – 4.8889 AI – 5.8148	Traditional – 1.78311 AI – 1.17791	0.029
Couch	Traditional – 4.1111 AI – 5.8148	Traditional – 1.57708 AI – 1.33119	<0.001
Lamp	Traditional – 4.0370 AI – 5.8889	Traditional – 1.62863 AI – 1.08604	<0.001
Plates	Traditional – 2.7037 AI – 4.8889	Traditional – 1.53960 AI – 1.62512	<0.001
Robe	Traditional – 4.0741 AI – 3.4074	Traditional – 1.51723 AI – 1.64689	0.128
Tablet	Traditional – 4.7407 AI – 3.8889	Traditional – 1.40309 AI – 1.47631	0.034

Table 6.2: **Perceived Trustworthiness** of products; average mean, standard deviation, and p-value

Product	Mean	Standard Deviation	P-Value
Shoe	Traditional – 5.0000 AI – 5.0000	Traditional – 1.24035 AI – 1.54422	1.000
Backpack	Traditional – 4.9630 AI – 5.4074	Traditional – 1.62863 AI – 1.67008	0.327
Couch	Traditional – 4.1111 AI – 4.8889	Traditional – 1.60128 AI – 1.57708	0.078
Lamp	Traditional – 3.7037 AI – 5.4444	Traditional – 1.48880 AI – 1.21950	<0.001
Plates	Traditional – 3.3333 AI – 5.1111	Traditional – 1.70970 AI – 1.39596	<0.001
Robe	Traditional – 4.0370 AI – 3.3704	Traditional – 1.80771 AI – 1.90441	0.193
Tablet	Traditional – 4.6667 AI – 3.6296	Traditional – 1.49358 AI – 1.41824	0.012

Table 6.3: **Purchase Intentions** of products; average mean, standard deviation, and p-value

Product	Mean	Standard Deviation	P-value
Shoe	Traditional – 4.4074 AI – 3.7778	Traditional – 1.88637 AI – 2.35884	0.284
Backpack	Traditional – 4.2963 AI – 5.0000	Traditional – 2.10886 AI – 1.73205	0.186
Couch	Traditional – 3.7037 AI – 4.7037	Traditional – 1.83586 AI – 1.89767	0.054
Lamp	Traditional – 2.7037 AI – 4.8889	Traditional – 1.38160 AI – 2.04438	<0.001
Plates	Traditional – 2.5926 AI – 4.1852	Traditional – 1.36605 AI – 1.59415	<0.01
Robe	Traditional – 3.5556 AI – 3.4444	Traditional – 1.88788 AI – 1.96769	0.833
Tablet	Traditional – 4.3704 AI – 3.4074	Traditional – 1.77911 AI – 1.50024	0.036

Table 7: **Bar Graph** of products; *Visual Appeal*

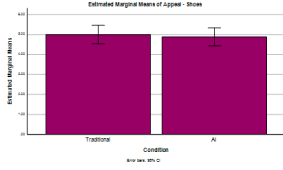
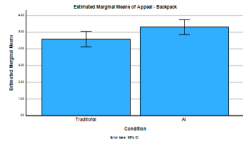
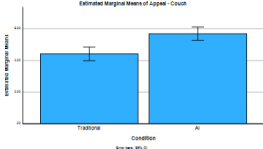
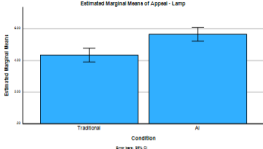
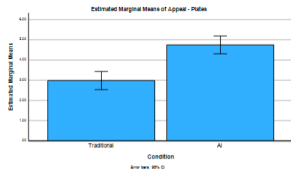
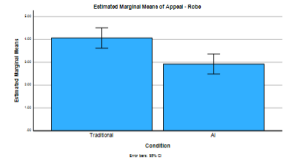
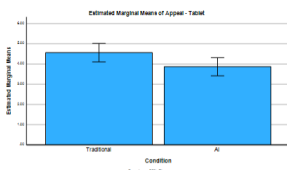
Figure 4.1 – Shoes	
Figure 4.2 – Backpack	
Figure 4.3 – Couch	
Figure 4.4 - Lamp	
Figure 4.5 – Plates	
Figure 4.6 – Robe	
Figure 4.7 – Tablet	

Table 8: **Bar Graph** of products; Perceived Trustworthiness

Table 5.1 – Shoes	Estimated Marginal Means of Trust - Shoes <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.6</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.6
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.6						
Table 5.2 – Backpack	Estimated Marginal Means of Trust - Backpack <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.9</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.9
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.9						
Table 5.3 – Couch	Estimated Marginal Means of Trust - Couch <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.9</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.9
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.9						
Table 5.4 – Lamp	Estimated Marginal Means of Trust - Lamp <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.9</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.9
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.9						
Table 5.5 – Plates	Estimated Marginal Means of Trust - Plates <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.9</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.9
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.9						
Table 5.6 – Robe	Estimated Marginal Means of Trust - Robe <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.6</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.6
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.6						
Table 5.7 – Tablet	Estimated Marginal Means of Trust - Tablet <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~3.8</td> </tr> <tr> <td>AI</td> <td>~3.6</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~3.8	AI	~3.6
Condition	Estimated Marginal Means						
Traditional	~3.8						
AI	~3.6						

Table 9: **Bar Graph** of products; Purchase Intentions

Figure 6.1 – Shoes	Estimated Marginal Means of Purchase - Shoes <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.7</td> </tr> <tr> <td>AI</td> <td>~1.6</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.7	AI	~1.6
Condition	Estimated Marginal Means						
Traditional	~1.7						
AI	~1.6						
Figure 6.2 - Backpack	Estimated Marginal Means of Purchase - Backpack <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.6</td> </tr> <tr> <td>AI</td> <td>~1.7</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.6	AI	~1.7
Condition	Estimated Marginal Means						
Traditional	~1.6						
AI	~1.7						
Figure 6.3 – Couch	Estimated Marginal Means of Purchase - Couch <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.6</td> </tr> <tr> <td>AI</td> <td>~1.8</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.6	AI	~1.8
Condition	Estimated Marginal Means						
Traditional	~1.6						
AI	~1.8						
Figure 6.4 – Lamp	Estimated Marginal Means of Purchase - Lamp <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.5</td> </tr> <tr> <td>AI</td> <td>~1.9</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.5	AI	~1.9
Condition	Estimated Marginal Means						
Traditional	~1.5						
AI	~1.9						
Figure 6.5 – Plates	Estimated Marginal Means of Purchase - Plates <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.5</td> </tr> <tr> <td>AI</td> <td>~1.8</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.5	AI	~1.8
Condition	Estimated Marginal Means						
Traditional	~1.5						
AI	~1.8						
Figure 6. 6 – Robe	Estimated Marginal Means of Purchase - Robe <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.6</td> </tr> <tr> <td>AI</td> <td>~1.5</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.6	AI	~1.5
Condition	Estimated Marginal Means						
Traditional	~1.6						
AI	~1.5						
Figure 6.7 - Tablet	Estimated Marginal Means of Purchase - Tablet <table border="1"> <thead> <tr> <th>Condition</th> <th>Estimated Marginal Means</th> </tr> </thead> <tbody> <tr> <td>Traditional</td> <td>~1.7</td> </tr> <tr> <td>AI</td> <td>~1.5</td> </tr> </tbody> </table>	Condition	Estimated Marginal Means	Traditional	~1.7	AI	~1.5
Condition	Estimated Marginal Means						
Traditional	~1.7						
AI	~1.5						

RESULTS

To conclude, this study aimed to explore the influence of AI-generated imagery on e-commerce consumer behavior, particularly in comparison with traditional imagery. Through an investigation of visual appeal, perceived trustworthiness, and purchase intention, the findings reveal several key insights.

Firstly, the introduction of AI-generated imagery does not necessarily diminish visual appeal; rather, its impact varies depending on the context and consumer preferences. Secondly, while traditional imagery may evoke a sense of familiarity and trust, AI-generated imagery has the potential to offer novelty and creativity, thereby influencing perceived trustworthiness in unique ways. This study highlights the interconnectedness of visual appeal, perceived trustworthiness, and purchase intention. It was found that visual appeal and perceived trustworthiness play pivotal roles, but it was not clear how purchase intent influences a consumer's decision to “buy” a product. The hypothesis was rejected because visual appeal and perceived trustworthiness increased for most products while purchase intentions remained unclear.

The AI products with the highest mean rating in visual appeal and perceived trustworthiness were the sofa, lamp, and plates. All had a p-value that was statistically significant (<0.001). The lamp

had the highest visual appeal with a mean rating of 5.89, the chair had a mean rating of 5.81, and the plates had a mean rating of 4.8.

Another conclusion with this study was that images with no human characteristics were more visually appealing and trustworthy compared to images with human characteristics (i.e. Backpack, Shoes, Tablet). These findings have significant implications for e-commerce businesses, suggesting that the integration of AI-generated imagery can offer a valuable avenue for enhancing consumer engagement and driving sales. By leveraging the strengths of both traditional and AI-generated imagery, businesses can tailor their visual content to resonate with diverse consumer preferences and foster trust in their brand. This study can also help AI companies to improve their generative AI models especially regarding improving images with human characteristics.

DISCUSSION

It's important to acknowledge the limitations of our study, such as the sample size, only 107 participants were screened. Although it was not a low number, more is needed to be concluded regarding purchase intent.

Future research could explore additional factors influencing consumer behavior in e-commerce, such as product type, brand reputation, and user-generated content.

REFERENCES

- Bledsoe, J. (n.d.). *Ecommerce sales & size forecast*. International Trade Administration | Trade.gov. <https://www.trade.gov/ecommerce-sales-size-forecast>
- Cropley, D. (2023). Generative AI and creativity: How chatgpt can enhance student creativity. *EDULEARN Proceedings*. <https://doi.org/10.21125/edulearn.2023.0253>
- Horton Jr, C. B., White, M. W., & Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-45202-3>
- Jrad, M. (2023). A role of artificial intelligence in the context of economy: Bibliometric analysis and systematic literature review. *International Journal of Membrane Science and Technology*, 10(3), 1563-1586. <https://doi.org/10.15379/ijmst.v10i3.1756>
- Kleinhenz, J. (n.d.). *Retail*. U.S. Chamber of Commerce | U.S. Chamber of Commerce. <https://www.uschamber.com/economy/2021-state-of-american-business-industry-perspectives/retail>
- Knight, S., & Bart, Y. (2023). Generative AI and user-generated content: Evidence from online reviews. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4621982>
- Nyaaba, M. (2023). Comparing human and AI's (GPT-4 and Gemini) understanding of the nature of science. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4661602>
- Reshaping E-Commerce: The Influence Of AI-Generated Content*. (2023, August 2). Forbes. <https://www.forbes.com/sites/shoplazza/2023/08/02/reshaping-e-commerce-the-influence-of-ai-generated-content/?sh=19b7982c273e>
- Su, Y., Zhang, K., Wang, Q., & Qiu, L. (2023). Generative AI and human knowledge sharing: Evidence from a natural experiment. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4628786>
- Zhang, Y., & Gosline, R. (2023). People's perceptions (and bias) toward creative content generated by AI (chatgpt-4), human experts, and Human-AI collaboration. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4453958>

Implementation of a Convolutional Neural Network for Diagnosis of Multiple Sclerosis

Megha Ramachandran

ABSTRACT

The potential applications of deep learning to transcriptome-based disease diagnosis are limited by the lack of large, publicly available gene expression datasets. Multiple sclerosis (MS) is a complex disease to diagnose due to the variety of ways it can present itself, and the lack of large amounts of accessible data makes deep learning applications to MS difficult. The purpose of this project is to test the efficacy of transfer learning (pre-training a model on a large dataset for use on a smaller one) on a Convolutional Neural Network (CNN) for diagnosis of MS from gene expression samples. The project also aimed to evaluate the benefits of self-attention, a mechanism that enables CNNs to extract spatially distant features, and KEGG-based gene ordering, which arranges associated genes near each other in input images. Four CNNs were pre-trained on patient gene expression samples from The Cancer Genome Atlas. Then, each CNN was fine-tuned on an MS dataset to classify patients as either having MS or not having MS. The fine-tuned models without KEGG ordering, both with and without self-attention, obtained an Area Under the Curve (AUC) greater than 0.7. The fine-tuned models with KEGG ordering, regardless of the addition of self-attention obtained lower AUCs, slightly greater than 0.6. This suggests that CNNs have potential as supplemental diagnostic tools for multiple sclerosis, especially as data availability increases. Further research can investigate alternate gene ordering methods, class activation mapping for model explainability, and other implementations of self-attention.

INTRODUCTION

This project aimed to implement a convolutional neural network (CNN) to accurately diagnose multiple sclerosis from patients' gene expression data. Multiple sclerosis is an autoimmune disease that occurs due to destruction of the myelin sheath, the fatty covering that protects nerve cells. This deterioration leads to a range of problems, such as fatigue, motor issues, loss of speech, and impaired vision. If undiagnosed, multiple sclerosis can be debilitating and can worsen over time. However, early diagnosis of multiple sclerosis can slow disease progression, making early detection crucial.

In 2019, it was discovered that the prevalence of multiple sclerosis in the United States

is twice as high the previously accepted estimate, at close to 1 million people across the country (Wallin et al., 2019). This indicates a growing demand for accurate multiple sclerosis diagnostic tools. Current diagnostic methods, such as the 2017 McDonald criteria, either use invasive procedures such as spinal taps, rely on the existence of lesions on the brain and/or spinal cord, or require the patient to have already experienced a symptomatic attack in order to confirm a diagnosis. The integration of genetic information to diagnostic criteria, or the identification of genetic biomarkers, can provide a supplemental means of identifying multiple sclerosis in patients before symptoms set in.

Prior research, such as that conducted in the

2020 research paper *Integrated single cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis*, has shown that multiple sclerosis leads to distinct gene expression signatures in blood cells (Schafflick et al., 2020). This indicates that transcriptome analysis is a viable means of supporting a multiple sclerosis diagnosis.

Machine learning methods have previously been applied to multiple sclerosis diagnosis. For example, the 2020 paper *Inferring Multiple Sclerosis Stages from the Blood Transcriptome via Machine Learning* used microarray data and shallow learning methods to successfully differentiate between patients with multiple sclerosis and healthy patients, as well as to discriminate between disease subtypes (Acquaviva et al., 2020). This paper showed high discriminative power between healthy and MS patients, as evinced by an accuracy of 88.7%, providing further support for gene expression as a multiple sclerosis diagnostic tool.

The 2018 paper *Gene Selection and Classification of Microarray Data Using Convolutional Neural Network* also applied machine learning methods, in this case, deep learning, to microarray data for various diseases, and was able to obtain a high accuracy for disease classification (Zeebaree et al., 2018). Therefore, neural networks have been shown to be effective at classifying disease stage from microarray data.

This project also aimed to investigate the implementation of self-attention, a mathematical mechanism that expands a CNN's ability to capture long-range spatial information, as it was implemented in the 2019 paper *Self-Attention Convolutional Neural Network for Improved MR Image Reconstruction* (Wu et al., 2019). The

workflow of the implemented self-attention mechanism is detailed in Figure 1, and is based off that used by Zhang et al., in the 2019 paper *Self-Attention Generative Adversarial Networks* (Zhang et al., 2019). It performs three 1x1 convolutions on the input feature map, and outputs an attention feature map with attention weights for each feature in the input. In the context of this project, it was believed that self-attention would allow the CNN to capture spatially-distant relationships between various genes.

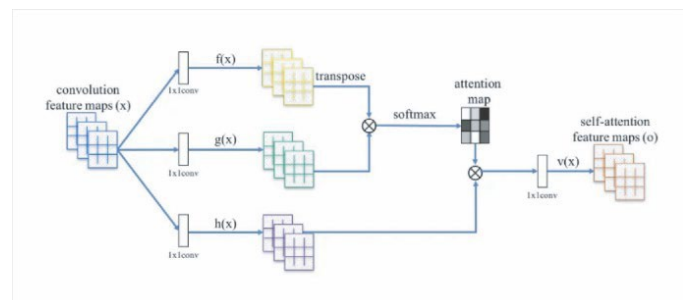


Figure 1. Diagram of self-attention mechanism as implemented by fast.ai and as used by Wu et al., 2019. Performs three 1x1 convolutions on feature map, then uses matrix multiplication to create an attention map. Outputs attention weights for each gene in the feature map.

Other notable methods used in this project were a transfer learning method to account for the limited multiple sclerosis dataset size, as well as the inclusion of KEGG gene ordering, as described in the 2020 paper *Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data* (López-García et al., 2020). Gene ordering methods are used to allow the CNN to extract logical relationships from genes that are placed near each other in input images, as these genes have related functions.

With the inclusion of self-attention, KEGG gene ordering, and transfer learning the goal of this

project was to obtain an Area Under the Curve of at least 0.7, as an indicator of model performance.

METHODOLOGY

Data Pre-Processing

3.1 Pre-training dataset (Cancer)

For transfer learning, a larger dataset was needed on which to pre-train the model. The dataset used was a publicly available dataset from The Cancer Genome Atlas (TCGA), which contained expression information for 10,535 patients across 33 cancer types, with expression information for >60,000 genes per patient. Based on the pre-processing methodology used by López-García et al., the dataset was separated into two classes, by those who either developed a new tumor event in a fixed time interval, and those who did not. Initial data dimensionality reduction was conducted using Mean Absolute Deviation (MAD) to reduce the total number of genes to 20,000. This identified the top 20,000 genes that showcased the most variable expression across patients. This dimensionality reduction is crucial to prevent model overfitting.

3.2 Fine-tuning dataset (Multiple Sclerosis)

For model fine-tuning, a publicly available gene expression dataset from the Gene Expression Omnibus (GEO) was downloaded, containing expression information for 815 patients, which is significantly smaller than the cancer dataset. To normalize the raw gene expression data, the dataset was first sorted by the coefficient of variation for each gene. The coefficient of variation is obtained by

dividing the standard deviation of the particular expression value by the mean expression value of that gene across all samples. Then, the data was z-scored to compress the values into a more manageable range.

Coefficient of Variation formula: $\frac{\sigma}{\mu}$

z-Score formula: $zz = \frac{xx - \mu}{\sigma}$

3.3 KEGG Gene Ordering

In the 2020 paper by López-García et al., a KEGG gene ordering method was used. First, a database of KEGG functional information was downloaded. This KEGG database contained information about the specific cellular processes associated with each human gene in the dataset. A mapping procedure was then performed, in which each gene in a dataset was associated to its KEGG functional information. This mapping procedure is outlined in the following diagram from López-García et al., and was applied to both the cancer and multiple sclerosis datasets.

Once this mapping procedure was performed, the next step was to arrange the genes in a 2-dimensional array, with genes that fall within the same KEGG functional subgroup placed near each other. The authors of the paper by López-García et al. utilized a treemapping approach, in which a hierarchical series of nested rectangles was created to group genes by their KEGG functions. The first step in this procedure was to create a treemap, as pictured below (Figure 2).



Figure 2. Visual representation of treemap from López-García et al., in which genes are organized by their KEGG functional information. The hierarchical structure of the treemap creates increasingly specific subdivisions regarding the gene's function.

At this point, this project's methodology differed from those employed by López-García et al., From this visual representation of the treemap, in which each pixel corresponds to a different gene, the locations of each gene in the treemap were then extracted as x- and y- coordinates. From these coordinates, a 2-dimensional 87 x 87 template gene array was created, which contained 7,0569 gene IDs, organized by their KEGG information. This template gene array is pictured below (Figure 3).

```
[ ['ENSG00000119121', 'ENSG00000099337'
  'ENSG00000006118', 'ENSG00000213920'
  ['ENSG00000070985', 'ENSG00000182255'
  'ENSG00000167693', 'ENSG00000080824'
  ['ENSG00000108556', 'ENSG00000165125'
  'ENSG00000108819', 'ENSG00000164715'
```

Figure 3. Section of template gene array. Gene IDs are extracted from treemap and arranged in a 2-dimensional array according to their location in the treemap.

Finally, individual patient gene expression values were mapping to the corresponding gene IDs in this template gene array, creating individual

patient gene expression arrays. A visualization of a patient gene expression array is shown below (Figure 4) in which each square corresponds to a gene, and the intensity of the square's color is representative of the intensity of the gene's expression.

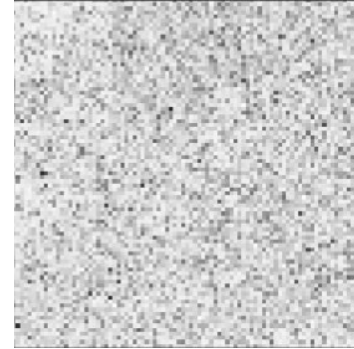


Figure 4. Visualization of a patient gene expression array from the multiple sclerosis dataset. Array dimensions are 87 x 87, with expression values for 7,5609 genes included. Intensity of expression value corresponds to the color of the associated pixel.

3.4 Data Oversampling

Both datasets had severe class imbalance, with far more samples of the positive class (new tumor event in the case of the cancer dataset, and presence of disease in the multiple sclerosis dataset) than the number of samples in the negative class. To prevent the model from classifying every sample as a member of the majority class and obtaining a falsely high accuracy, oversampling was necessary. First, a train/test split was conducted on both datasets. For the cancer dataset, a train/test split of 90/10 was used. For the multiple sclerosis dataset, a train/test split of 70/30 was used. Then, samples in minority class in each of the train/test sets was duplicated to obtain relatively equal numbers of samples in both classes.

Model Creation and Evaluation

3.5 Model Architecture

Multiple models were tested with variations of hyperparameters such as the network depth, size of convolutional and pooling kernels, and the number of channels. To minimize the number of trainable parameters while still retaining model accuracy, the following model architecture was decided upon, comprised of two convolutional blocks, the self-attention mechanism, a 1x1 convolutional layer, and the fully connected layer.

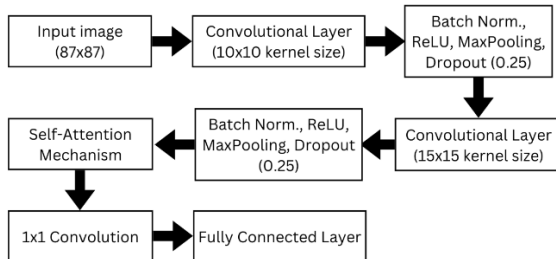


Figure 5. Architecture of convolutional neural network.

3.6 Model Training

Four versions of the neural network were tested:

- 1) With self-attention, with KEGG gene ordering
- 2) Without self-attention, with KEGG gene ordering
- 3) With self-attention, without KEGG gene ordering
- 4) Without self-attention, without KEGG gene ordering

This was done to observe the effects of self-attention and KEGG gene ordering. For the models without KEGG gene ordering, the genes in the input images were organized in a random order. The entire model was first pre-trained on the cancer dataset. Then, the model weights were saved and applied to

the multiple sclerosis dataset, and certain layers were re-trained.

If the model had self-attention, the following layers were re-trained on the multiple sclerosis dataset:

- 1) Self-attention mechanism and its associated layers
- 2) 1x1 convolutional layer
- 3) Fully connected layer

For the models without self-attention, only the last fully connected layer was re-trained when applied to the multiple sclerosis data. Area Under the Curve of the Receiver Operating Curve (AUC-ROC), which is a measure of the model's classification ability, was calculated for model performance on both the cancer and multiple sclerosis datasets.

RESULTS

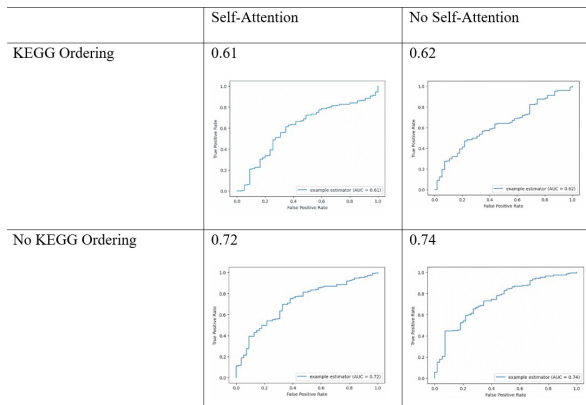
4.1 Cancer dataset

First, model performance after pre-training on the cancer dataset was evaluated. Area Under the Curve values are recorded in the below data table.

	Self-Attention	No Self-Attention
KEGG Ordering	0.74	0.72
No KEGG Ordering	0.70	0.69

4.2 Multiple Sclerosis dataset

Area Under the Curve values for model performance after fine-tuning on multiple sclerosis data, as well as Area Under the Curve graphs, are recorded in the below data table.



CONCLUSION

5.1 Cancer dataset

The model with KEGG ordering and self-attention obtained the highest Area Under the Curve value, indicating the highest discriminative power between the classes. Firstly, this indicates that, for the cancer dataset, the implementation of self-attention may have contributed to an improved model accuracy, although more trials would have to be conducted to determine if the improvement was significant. Secondly, KEGG gene ordering did appear to improve model performance, as it led to a slightly improved AUC. This indicates the potential efficacy of gene ordering in helping the CNN extract meaningful features.

5.2 Multiple sclerosis dataset

Both models without KEGG ordering obtained AUCs greater than 0.7, the model without KEGG ordering obtaining the highest AUC of 0.74. The models with KEGG ordering had significantly lower performance, with AUCs slightly greater than 0.6. This suggests that the relationships learned from the KEGG ordering in the cancer dataset were not applicable to the multiple sclerosis dataset. The effects of this could be mitigated in the future by re-

training more layers of the network on the multiple sclerosis dataset.

Additionally, self-attention did not appear to improve model performance. This may have been because the number of samples in the multiple sclerosis dataset was not high enough to effectively re-train the self-attention layer.

DISCUSSION

Both initial engineering goals were met. An AUC of at least 0.7 on the multiple sclerosis dataset was obtained, and self-attention was successfully implemented within the model architecture. This project was novel in its application of a convolutional neural network to gene expression data for multiple sclerosis, as well the integration of self-attention, gene ordering, and transfer learning. The model's performance on the multiple sclerosis dataset indicates that convolutional neural networks are viable tools for disease diagnosis from microarray data. Microarray data comes from patient blood samples, which is a less invasive procedure than current diagnostic methods such as spinal taps and could be incorporated more easily into a clinical routine. Additionally, the results indicate the potential efficacy of techniques such as gene ordering and transfer learning. However, these results did not outperform existing results from shallow learning techniques, such as those obtained in the paper by Acquaviva et al., 2020.

Despite this, convolutional neural networks techniques like this may present added benefits due to their usage of unsupervised feature selection. Once higher model performance is obtained, model explainability techniques such as Class Activation

Mapping could be applied to as a means of extracting a list of candidate genetic biomarkers that are used by the network to classify the samples. Future research can also include the investigation of other CNN model architectures, further

dimensionality reduction of the dataset, decreasing model parameters to enable better generalization, and lastly, extending the model to multi-class classification to differentiate between subtypes of multiple sclerosis.

REFERENCES

- Acquaviva, M., Menon, R., di Dario, M., & Dalla Costa, G. (2020, July 21). Inferring multiple sclerosis stages from the blood transcriptome via machine learning. *Science Direct*.
<https://www.sciencedirect.com/science/article/pii/S2666379120300690>
- López-García, G., Jerez, J. M., Franco, L., & Veredas, F. J. (2020). Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data. *PLOS ONE*, 15(3).
<https://doi.org/10.1371/journal.pone.0230536>
- Schafflick, D., Xu, C. A., Hartlehnert, M., Cole, M., Schulte-Mecklenbeck, A., Lautwein, T., Wolbert, J., Heming, M., Meuth, S. G., Kuhlmann, T., Gross, C. C., Wiendl, H., Yosef, N., & Meyer Zu Horste, G. (2020, January 14). Integrated single cell analysis of Blood and Cerebrospinal Fluid Leukocytes in multiple sclerosis. *Nature communications*.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6959356>
- Wu, Y., Ma, Y., Liu, J., Du, J., & Xing, L. (2019). Self-Attention Convolutional Neural Network for Improved MR Image Reconstruction. *Information sciences*, 490, 317–328.
<https://doi.org/10.1016/j.ins.2019.03.080>
- Zeebaree, D. Q., Haron, H., & Abdulazeez, A. M. (2018). Gene selection and classification of microarray data using Convolutional Neural Network. *2018 International Conference on Advanced Science and Engineering (ICOASE)*. <https://doi.org/10.1109/icoase.2018.8548836>
- Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2019). Self-Attention Generative Adversarial Networks. <https://arxiv.org/abs/1805.08318>

Stock Wizard: A Machine Learning Model Designed for Day Trading Stock Price Prediction

Ryan Chung

ABSTRACT:

The stock market comes with the risk of monetary loss, but what if there was a way to predict the trends and fluctuations of daily stock prices for a higher chance of guaranteeing positive returns? Day-trading as an investment strategy is currently riddled with risk, and when executed improperly, is no different than betting money on a coin flip. This project will reduce the risk of day-trading volatile stocks and maximize returns through a machine learning prediction model. Specifically, this project uses eXtreme Gradient Boosting, a relatively new, regression-based Decision Tree machine learning technique. The datapoints for training were daily oil and stock index prices (Nasdaq, NYSE, and S & P 500) in addition to 5-minute intervals of volume and high/low prices for each company (Google, Amazon, Visa, Walmart, and Exxon Mobil). After training the model with 8000 data points from each of the 5 companies, the model was able to obtain a 92% accuracy at predicting whether or not a stock would go up or down, an 80¢ average error when predicting the exact price of the stock, and a median percentile error of 54¢. Overall, this project offers strong evidence that eXtreme Gradient Boosting is a capable, accurate, and viable machine learning model for day-trading stock price prediction. With additional data, this project will be able to forecast company share prices year-round and with a slimmer margin of error, as the current results are only based on 5 months of data.

INTRODUCTION:

Through the use of AI and Machine Learning Techniques, this project will offer evidence that stock day trading trends are predictable and can earn a guaranteed profit by creating a machine learning model to predict the stock prices of major U.S. companies every 5 minutes.

Stocks are defined as individual shares of a company, which are reflective of a company's profit and success in their industry. When a company decides they want to sell stocks, they divide their company into shares, which are each worth a certain amount of money. The initial price of the stock is the Initial Public Offering, or IPO. Based on demand, the stock price will then go up or down. If the company is successful, there is high demand, which

in turn, causes prices to rise. If the stock is mainly being sold though, the price drops. Often times, investors will sell/buy based on economic cues and company finance reports. The expectation is that if the company earns a higher profit, then its stock price will go up. Shareholders and investors, if buying stocks from the correct companies, should earn money too, as they are investing in specific companies and banking on the stock price to rise, allowing them to earn money when they sell.

While the premise is extremely simple, there are many ways investors manipulate the stock market. The vast majority of the population buys stocks and just holds on to them. Almost all stocks, over time, will increase in price, so people are taught to play the long game. Even if the stock initially

goes down, shareholders bank on the fact that the price will rebound. This is a decently safe way to invest in the stock market, but it offers only modest returns after a long period of waiting. Due to this, some people turn to day-trading. Stock prices, throughout the day, fluctuate constantly. Therefore, it is possible to buy during the day with an exorbitant amount of money, and a minute later, when the stock goes up slightly, sell. This high-risk, high-reward method of investing offers the best returns at a rapid rate, albeit with much more added risk. Often times, this risk of monetary loss causes people to stray away from the stock market, even avoiding the fairly safe plan of long-term stock investing.

The statistics on monetary loss in the stock market are quite shocking. Americans, in 2022, lost 9 trillion dollars from the stock market according to the Federal Reserve, or equivalent to 36 companies worth as much as Apple (Frank, 2022). The total monetary value of the stock market fell from 42 trillion dollars to 33 billion. According to Mark Zandi, an economist at CNBC, these losses could hurt real GDP growth by 0.2 percent, or around 40 billion dollars (Zandi, 2022). Average Americans, or Americans at median income, the Federal Reserve also tells us, have in total lost 70 billion dollars in stock wealth. This decrease in wealth and economic health has a positive feedback loop, where decreased consumer spending and lower stock prices will drive down the economy even further. Add this on to the recent interest rate increases and record inflation, and the U.S. economy is in a precarious state.

Part of the reason for these massive losses in the stock market is its volatility. The variety of factors affecting the stock market have made it almost impossible to predict. With that said, the

astronomical rise of AI in our current time could change this. The rise of the chatbot Chat GPT has already shocked the world with its abilities. AI's future is currently wide open, and there is no prediction of what it can do in the future. A prediction tool using modern computing technology would help the U.S. by giving the government and investors a model to inform their financial decisions. Currently, there are other machine learning models that can predict the closing prices of stocks over an extended period of time with decent accuracy. For example, in *Machine Learning for Stock Prediction Based on Fundamental Analysis*, Yuxuan Huang, in 2021, was able to create an ML model based on the Random Forest network that earned mean quarterly relative return of 1.63%, compared to the universal return of 0.0164% (Huang et al, 2021). In *Machine Learning Techniques for Cross-Sectional Equity Returns Prediction*, Christian Fieberg was able to employ a Support Vector Regression ML model that obtained the low R Squared value of 0.34 (Fieberg et al, 2020). That said though, there have almost been no noteworthy attempts at predicting day-trading trends and stock prices within a smaller timespan. Day trading offers quite a few benefits compared to long-term stock prediction, including faster returns, and potentially more returns, but it entails much more risk of loss and requires a larger sum of money to make a meaningful profit. A day trading predictive machine learning model would greatly decrease the risk of loss in day trading and would help maximize revenue. In addition, this project will be adding many other data points like interest rates, volume, and company job numbers to a gradient boosting machine learning model to increase

accuracy, which has not been done on other stock predicting models.

Looking at previous work in this subject, there seems to be no consensus on one Machine Learning Technique. Some articles state that Random Forest is the most accurate ML technique (Huang, 2021), while other state that ADA Boost reigns supreme (Khan, 2023). Indeed, there are several great ML programs that would be proficient at predicting stock prices, but in the end, this project decided to implement the Extreme Gradient Boosting Machine Learning Model, or xGBoost. The biggest reason for settling on this technique was the lack of articles on predicting stock prices with xGBoost. Random Forest, Central Neural Network, and Support Vector Regression all had already been used to predict stock values several times, but xGBoost only had two instances of previous experimentation. This study would be another step towards uncovering what xGBoost's true talent is. Adding on, Extreme Gradient Boosting has always been considered excellent at predicting values throughout time and finding trends in linear data, meaning it should be a great option. Overall, it is hypothesized that this ML model will be able to predict the stock price accurately and earn a return of over 10.26%, which has been the average return from the S & P 500 over the past 60 years.

MATERIALS AND PROCEDURE

Before the model was coded, data and materials had to be compiled. This model was run coded and run on a Lenovo Laptop with an Intel I5 Processor using Microsoft Visual Studio Code. Data wise, the basic stock statistics are necessary, including the stock's opening price, category, high, low, closing price, and volume every 5 minutes. All

of these values were obtained from Stooq.com. In addition to this, the daily prices of the S&P 500 and the NASDAQ stock index were obtained from Yahoofinance.com. Finally, daily oil prices around the U.S. were obtained from Marketwatch.com. Collectively, the data should present an accurate picture of the economy's path in addition to the minute daily trends each stock undergoes. This data will all be compiled from the time span of September, October, November, December, January, and February of 2023 to 2024.

For data pre-processing Microsoft Visual Studio was downloaded from the Microsoft Store. After that, Python was downloaded and imported to Visual Studio. The Math Plot Library was also imported, as it is able to graph out our predicted and expected stock prices. After all the data was downloaded, the program iterated through all the data points and checked for missing values. Any absent values in the data were deleted. In addition, all outliers in the dataset, defined as values which are more than 20\$ off their neighbors, were removed. After the data was cleaned, one final check was conducted to ensure the data was copied properly. With the data properly organized and cleaned, the Math Plot Library was then used, which printed out the expected vs predicted graphs. To ensure the code was working, a preliminary check for the line graph. If the line graph matches, the ML Model will proceed to part 2 of the procedure.

Stock Price over Time

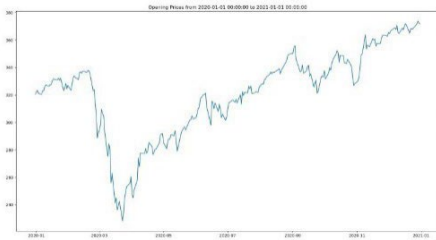


Figure 1: When output from the code, the line graph formatting and layout should look like this. This graph specifically compares the closing price of a stock over the period of a month.

The second part of the procedure is the creation of the Machine Learning Model. For this project, Extreme Gradient Boosting Regression was used. Simply put, it is a combination of weak learning models, eventually calculating an accurate value through the teamwork of simple computer decision tree models. In this specific model, the process is no different. First, the xGBoost model will require data. For the intent of explanation, the dataset will be small and simple, as shown below. Our real data set will contain thousands of lines and 6-7 factors.

Amazon and Google Stock Data Table

Stock Name	Price	Volume	Date
Amazon	113	58.7 Million	11/23/23, 3:55 P.M
Amazon	156	48.6 million	4/23/23, 4:55 P.M.
Google	203	102 million	1/23/22, 8:55 A.M.
Amazon	90	30.4 million	9/11/22, 12:55 P.M.
Google	198	78.7 million	6/18/23, 4:45 P.M.

Table 1: This data table contains our working data set. It has two different companies, 5 stocks total, with each stock's price, volume and time listed.

Once the ML model has received the data, its first guess, if it is trying to predict the price, will be the average of all the prices, which is 152. After that, the model will calculate its error. The difference is calculated between each stock's actual price and

the predicted price. The difference, called a pseudo-residual, is stored along with each corresponding stock. For the next step, the model actual pivots and starts trying to predict the pseudo-residual. This is used through small "leaves", which make up each decision tree, like the one shown below.

Gradient Boosting Regression "Leaf" Example

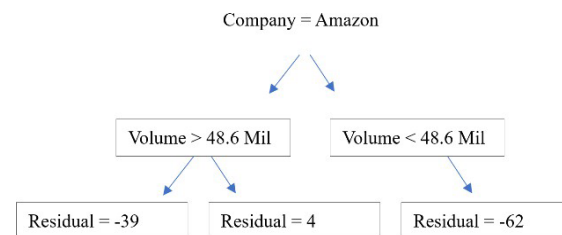


Figure 2: This is a simplified flowchart of the decision-making Gradient Boosting does. First, the model starts with the company name of the stock. Then, based on the volume, we can further classify the stock price, eventually calculating the residual. This is only one of the "leaves" on a full tree of them.

Now, xGBoost combines the model we just made with the first decision tree that we created. By adding the residuals to the first predicted price, the model obtains its second predicted price. In this case, the predicted value for the Amazon Stock in the first row of the table will be 113, which is actually exactly the same as my expected data value. This is too accurate though. The model will need additional tweaking, as the current model can only predict values for these specific data points right now. xGBoost solves this problem of overfitting by multiplying every residual with a learning rate, which is typically a value between 0 and 0.1. This scales down the impact each residual has on the final model and ensures that it is adaptable to different datasets. Again, looking back out our first data point, if I multiply the first against the manipulated residual my prediction would now be

$152 + (0.1 * -39)$, which equals 148.1. This prediction doesn't seem as accurate, but it is better than the old model, which stated that every Amazon stock price would be 113 dollars when the volume was less than 48.6 million. This learning rate means that the prediction tree is a small step in the right direction.

From now on, the model repeats this process. It continually creates new "leaves", which it adds to its decision tree. In this case, the 3rd guess from this model would be the original price subtracted by the new residual. These baby steps are repeated a fixed number of times, and eventually, the model is able to create a decision tree that adds/subtracts a certain value based on its factors to give an accurate predicted value. Another visual of this process is included below.

Gradient Boosting Regression Model

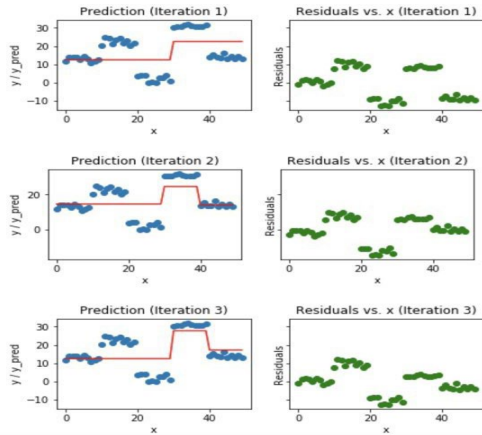


Figure 3: Starting from the top, the three graphs are in chronological order, depicting three sample steps of Gradient Boosting Regression. On the left, the blue dots are the data points that need to be predicted. The red line is the value our machine learning model would predict after that round of training. The graphs on the right show the error rate, or the residuals of each prediction. As the machine learning models continue to iterate/improve, the trend line starts to become more and more accurate while the residuals are getting closer and closer.

After the Model has been trained, the output data values will be analyzed and checked for accuracy, which will be explained in the data analysis section.

DATA ANALYSIS

There are three standard measurements for accuracy testing and statistical analysis among the stock prediction machine learning models. These equations are Standard Mean Absolute Percentage Error, Percentile Error, and Root Means Squared Error. The equations for each of the statistics are as follows:

Standard Mean Absolute Percentage Error (SMAPE)

$$\text{SMAPE} = \frac{1}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{(A_t + F_t)/2}$$

N = number of data points, A_t = expected price, F_t = predicted price.

Calculates the average percentage error of the xGBoost Model.

Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$

N = number of data points, y_j = expected price, $\hat{y}_j$ = predicted price.

Calculates average dollar error of the ML Model.

50% Percentile Error

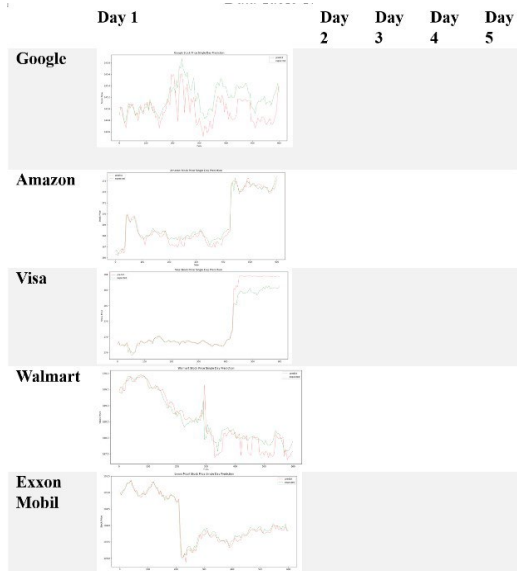
$$\text{error \%} = \left(\frac{A - B}{A} \right) \times 100\%$$

A = Expected Value, B = Predicted Value.

After the percentage error is calculated for every single predicted vs. expected value, the percentage error is ordered from least to greatest, and the median is taken.

All of the data will be stored in Data Table 2 and graphed out for visual and numerical evidence of accuracy. To evaluate the overall performance of our model, the statistical results from each individual company will be averaged.

RESULTS:



Data Table 1:

Table 3: This table contains the line graphs of the stock prices for visual analysis of the expected vs. predicted trends and the accuracy of the model.

Advanced Statistical Accuracy Analysis

	SMAPE	RMSE	% Error
Google	0.270878	0.157190	0.232503
Visa	0.986920	0.208926	0.123375
Walmart	0.271678	0.108760	0.134371
Amazon	0.873996	0.385702	0.402769
Exxon Mobil	0.081296	0.058595	0.051190

Table 2: This Data Table stores the data values of each statistical analysis measure.

CONCLUSION:

In the end, the machine learning model was able to obtain a RMSE of 80 cents, a 50% percentile error of 54 cents, and a 92% accuracy at determining whether or not a stock would go up or down. The 92% accuracy at predicting the overall trend of a stock price is an extremely reassuring predictor of

success for this ML Model. A stock price predictor certainly attempts to be as accurate as possible, but as long as money is being made, it can be considered a success. That said though, the xGBoost model also displays tendencies of high accuracy. An RMSE of 80 cents showcases that the model is highly accurate, as the average error is less than a dollar. To put it simply, the stock prediction model, even with stock prices in the hundreds, was almost always within 80 cents of the actual price. In consideration of the current scope of potentially several days, an 80-cent error rate is a small, negligible difference. Likewise, a 50% percentile error of 54 cents is very encouraging, as it shows that this model, without outliers, is accurate to almost a half a dollar. Again, this level of precision has not been seen in many of the other ML Models online. Since the stocks used to test this model had prices in the \$100 to \$200 range, the percentage error is less than 1%, which is quite impressive.

This high accuracy also offers evidence that this prediction model will be accurate in the future, as the ML model has room for error. The fairly primitive data used while training this model also offers plenty of room for improvement. Additional data, in addition to other types of data, as mentioned below in the “Future Research” Section, could be added. The engineering goals for this project were met, as a 10.26% return was surpassed, with a 39% return scenario being the best-case prediction from the model. This was obtained from buying and selling Exxon Mobil selectively during the 5 days of testing throughout specific peaks and falls in the day-to-day stock graph. While the stock trends and prediction scope were limited to 24 hours, the buying and selling was done on an hourly basis.

FUTURE RESEARCH:

For future research, more data from every quarter of the year is necessary, and attempts will need to be made to remove fluctuations within the trend charts and accuracy during sudden changes in stock price. This can be done by incorporating more data and tuning the model to fit the parameters more. Examples of data that could be added to the model include P/E, projections from investors, and volatility. In addition, to further increase model scalability and adaptability, fundamental analysis, or the inclusion of expert testimony and news articles could be used in addition to the technical analysis already being implemented in the model. The overall precision of the model

would be increased, but most noticeably, the accuracy in instances of earnings reports, layoffs, and other major company news should skyrocket, as there is now information on the real-world state of the company. Finally, it will also be necessary to invest time into researching the optimal scope for the xGBoost model. The current time period used is 24 hours, but there is potential for increased returns if the scope of practice was shortened/lengthened

ACKNOWLEDGMENTS:

I would like to thank my mentor Dr. Robert Chen for all his assistance throughout the various stages of my project, especially when I struggled with model tuning. In addition, I would like to thank the academic peers that took the time to read through my research plan, my project methodology, and my research paper. Last but not least, I would like to thank Ms. Allender for teaching me how to navigate the CSRSEF and WSSEF gauntlet.

WORKS CITED:

- Kumar, Deepak, et al. "A Systematic Review of Stock Market Prediction Using Machine Learning and Statistical Techniques." *Materials Today: Proceedings*, Jan. 2021, <https://doi.org/10.1016/j.matpr.2020.11.399>.
- Kumbure, Mahinda Mailagaha, et al. "Machine Learning Techniques and Data for Stock Market Forecasting: A Literature Review." *Expert Systems with Applications*, vol. 197, July 2022, p. 116659, <https://doi.org/10.1016/j.eswa.2022.116659>.
- Nabi, Rebwar M., et al. "A Novel Approach for Stock Price Prediction Using Gradient Boosting Machine with Feature Engineering (GBM-WFE)." *Kurdistan Journal of Applied Research*, vol. 5, no. 1, 30 Apr. 2020, pp. 28–48, <https://doi.org/10.24017/science.2020.1.3>. Accessed 3 Dec. 2023.
- Paspanthong, Art, et al. *Machine Learning in Intraday Stock Trading*.
- Singh, Tinku, et al. "An Efficient Real-Time Stock Prediction Exploiting Incremental Learning and Deep Learning." *Evolving Systems*, vol. 14, no. 6, 21 Dec. 2022, pp. 919–937, <https://doi.org/10.1007/s12530-022-09481-x>. Accessed 3 Dec. 2023.
- Subasi, Abdulhamit, et al. "Stock Market Prediction Using Machine Learning." *Procedia Computer Science*, vol. 194, 2021, pp. 173–179, <https://doi.org/10.1016/j.procs.2021.10.071>.
- Wang, Bicheng, and Xinyi Zhang. *Deep Learning Applying on Stock Trading*.

Using Behavioral Forma Mentis Networks to Map Gender Bias in Large Language Models

Sarah Wen

ABSTRACT

Monitoring gender bias in Gemini (generative AI chatbot) is key to understanding how big of a problem implicit gender bias is in chatbots. This project seeks to investigate how *behavioral forma mentis networks (BFMN)*, which represent information as a web of concepts, can be used to analyze the context behind how Gemini displays gender bias and establish a benchmark of gender bias in the developing world of AI. By employing psycholinguistic tools to dissect Gemini's conceptual associations through free association (ex. “given the word ‘dog’, what is the first word you think of?”) and affect norms (more feminine or masculine?), this project illuminated how chatbots perceive careers. Through Google Colaboratory, Gemini API was prompted for the first three words that came to mind when given a cue word (‘nurse’, ‘assistant’, ‘professor’, ‘leader’, ‘lawyer’). Initial findings revealed that “feminine” roles (‘nurse’ and ‘assistant’) are characterized by nurturing traits (‘compassionate’ and ‘aid’), while “masculine” roles (‘lawyer’ and ‘professor’) are associated with strength (‘courage’ and ‘respect’). Extreme ends in the study exposed the magnitude of bias embedded within Gemini's responses with the BFMN giving cue word “assistant” = 4.1944 and “lawyer” = 1.8576 (out of 5, 1 = strongly masculine, 5 = strongly feminine). These insights not only shed light on the current state of framing gender bias in AI but also pave the way for leveraging cognitive networks to cultivate more unbiased generative AI systems in the future, prompting a critical reevaluation of data infrastructure used in training LLMs.

INTRODUCTION

According to the 2023 Gender Social Norms Index (GSNI) published by the United Nations Development Programme, almost 90% of men and women worldwide still hold bias against woman. When ChatGPT, a generative artificial intelligence chatbot, was released by OpenAI in November 2022, it became one of the fastest-growing consumer platforms ever, reaching 1 million users in five days and 100 million in less than three months. With the introduction of such generative AI platforms, there is growing concern about the potential for these AI systems to perpetuate and amplify implicit gender biases present (Vlasceanu & Amodio, 2023). If AI

systems generate content that influence users’ behavior and decisions, it can have significant implications for individuals and society. For example, in 2014, Amazon created an AI recruiting tool that rated job candidates scores based on their resumes, only to find that their male-dominated results reflected the male dominance in Big Tech.

As a result, there is a pressing need to understand and quantify in what context large language models (LLMs) frame implicit gender bias. A potential solution for this problem stems from a relatively new type of cognitive network modeling known as a behavioral forma mentis network (BFMN) which can be created to map associated

responses to memory recall through free association, which is a psychological concept that refers to the process of retrieving information from memory by allowing thoughts and ideas to flow freely without a structured effort. BFMNs can represent the given knowledge as a web of concepts rated by valence (pleasantness of a stimulus) and arousal (intensity of emotion provoked by stimulus) interconnected by the free association links. Research into *forma mentis networks* have not yet been applied to analyze how LLMs use contextual clues to “unconsciously” produce responses with gender bias. Ultimately, my research into BFMNs led to my research question of *“Can a Behavioral Forma Mentis Network provide context to how Gemini approaches the framing of gender bias regarding careers that are stereotypically male-dominated or female-dominated?”* Through a methodological approach encompassing data collection, processing, and analysis, this study aimed to contribute to a deeper understanding of the complex interplay between AI cognition and societal biases with Google Gemini, with implications for the development of more ethical and inclusive AI systems.

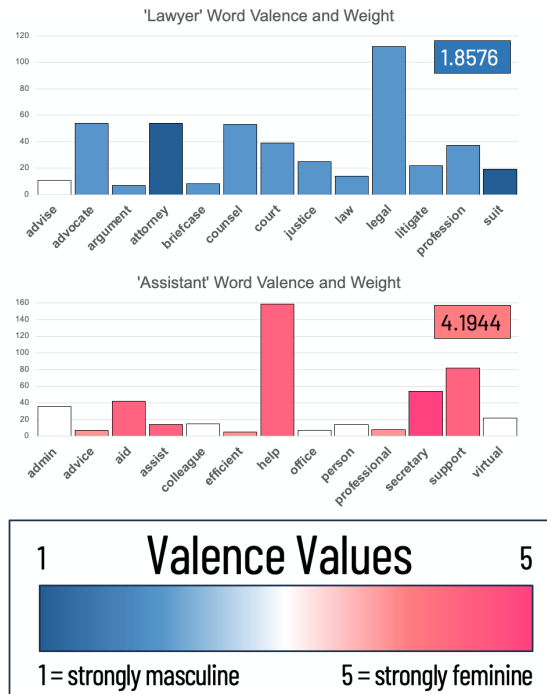
METHODOLOGY

My data collection methodology spanned three steps: free association, word stemming, and valence. First, I coded a program in Google Colaboratory (Python) that connected to the Gemini API that would allow direct interfacing with Gemini’s servers to automate the prompting process. With free association, the process where a word spontaneously conjures the idea of another word – BFMNs can give context to the “thinking” an LLM does when approaching a prompt. Through this code in Google Colaboratory, 169 trials were done for Gemini to

collect 3 words per trial, for all of the five careers (‘nurse’, ‘lawyer’, ‘professor’, ‘leader’, ‘assistant’), leading to a total count of 2535 words collected. In each trial, the Gemini model was given the prompt “given the word [insert career here] what are the first three words that come to mind”. At the end of the 169 trials, I set the program to automatically download the csv file, producing 1 for each of the 5 careers. In between the free association and word stemming, quality checks of the words themselves were manually done to ensure validity and accuracy. For example, I would find words in different languages or numbers instead of words, which needed to be cut out of the dataset, taking out that entire trial. Next, I coded a Python Google Colaboratory that implemented the Porter Stemmer NLP to execute word stemming. With word stemming, the generated words from the free association stage were converted into the root word to create precise datasets. As example, the words “joyful” and “joyous” would be converted into “joy” and if “joyful” showed up 4 times and “joyous” showed up 2, they would be combined into “joy” with a count of 6. While the Porter Stemmer did perform better compared to other existing NLP stemmers, words produced by the Porter Stemmer weren’t always real words. Therefore, to ensure that the words were accurately stemmed, I also manually stemmed all of the 2535 words and compared them to the Porter Stemmer. Lastly, another Python program was coded (and implemented the Gemini API) to identify the valence of all of the 2525 words collected. Typically, the valence is defined as emotional sentiment, as in “how positive or negative is this word”. However, here, I identified valence as “how masculine or feminine of a characteristic is

this word” on a scale of 1 to 5, where 1 was strongly masculine and 5 was strongly feminine. To analyze the data collection, I used valence and count of each

word in that dataset (the 5 careers) and took a weighted average to output the BFMN score of each career.



Resulting graphs for the five careers with the BFMN score labeled at the top of each graph

RESULTS

Initial findings revealed that “feminine” roles (‘nurse’ and ‘assistant’) are characterized by nurturing traits (‘compassionate’ and ‘help’), while “masculine” roles (‘lawyer’ and ‘leader’) are associated with strength traits (‘vision’ and ‘advocate’). Based on the data collected and resulting graphs produced, two conclusions could be made. First, the ‘lawyer’ graph is filled with light and dark blues, leading to the conclusion that Gemini’s thought process about lawyers runs through more masculine channels. On the other hand, the ‘assistant’ graph is pink, leading to the conclusion that Gemini’s thought process around ‘assistants’ runs through more feminine channels.

Indeed, the extreme ends in the study exposed the magnitude of bias embedded within Gemini’s responses with the BFMN giving cue word “assistant” = 4.1944 and “lawyer” = 1.8576 (of out 5). Lastly, I was able to identify common threads that gave crucial context to how Gemini approaches prompts regarding nurse, lawyers, leaders, assistants, and professors, achieving my engineering goal of understanding how Gemini displays implicit gender bias.

DISCUSSION

A future step for this experiment would be to test different careers that in the same industry (i.e. doctors and nurses), as this would allow for direct

comparison apples-to-apples comparison. As professions and their designated roles shift over time, analyzing these changes as they occur would allow for more accurate data to map the discrepancies in mindset. Another future step to take could be expanding the number of generative AI tools assessed to encompass the top 3 (Google

Gemini, OpenAI ChatGPT, Microsoft Copilot). Through the datasets used to the developers behind these applications, there will be more noticeable differences in their responses and tracking the comparisons will allow for impactful collaboration between these developers.

REFERENCES

- Abramski, K., Citraro, S., Lombardi, L., Rossetti, G., & Stella, M. (2023). Cognitive Network Science Reveals Bias in GPT-3, GPT-3.5 Turbo, and GPT-4 Mirroring Math Anxiety in High-School Students. *Big Data and Cognitive Computing*, 7(3), 124–124. <https://doi.org/10.3390/bdcc7030124>
- Brian, K., & Stella, M. (2023). Introducing mindset streams to investigate stances towards STEM in high school students and experts. *Physica A: Statistical Mechanics and Its Applications*, 626, 129074–129074. <https://doi.org/10.1016/j.physa.2023.129074>
- De Deyne, S., Navarro, D. J., & Storms, G. (2012). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods*, 45(2), 480–498. <https://doi.org/10.3758/s13428-012-0260-7>
- Hagendorff, T. (2023). Machine Psychology: Investigating Emergent Capabilities and Behavior in Large Language Models Using Psychological Methods. <https://arxiv.org/pdf/2303.13988.pdf>
- Hicks, M. (2021). Sexism Is a Feature, Not a Bug. In *Your Computer Is on Fire* (pp. 135–155). doi:10.7551/mitpress/10993.003.0011
- Liang, L., & Acuna, D. E. (2019). Artificial mental phenomena: Psychophysics as a framework to detect perception biases in AI models. *ArXiv:1912.10818 [Cs]*. <https://doi.org/10.1145/3351095.3375623>
- Stella, M. (2020). Forma Mentis Networks Reconstruct How Italian High Schoolers and International STEM Experts Perceive Teachers, Students, Scientists, and School. *Education Sciences*, 10(1), 17. <https://doi.org/10.3390/educsci10010017>
- Stella, M., de Nigris, S., Aloric, A., & Siew, C. S. Q. (2019). Forma mentis networks quantify crucial differences in STEM perception between students and experts. *PLOS ONE*, 14(10), e0222870. <https://doi.org/10.1371/journal.pone.0222870>
- Stella, M., Kapuza, A., Cramer, C., & Uzzo, S. (2020, July 18). Mapping computational thinking mindsets between educational levels with cognitive network science. *ArXiv.org*. <https://doi.org/10.48550/arXiv.2007.09402>
- Stella, M., Vitevitch, M. S., & Botta, F. (2022). Cognitive Networks Extract Insights on COVID-19 Vaccines from English and Italian Popular Tweets: Anticipation, Logistics, Conspiracy and Loss of Trust. *Big Data and Cognitive Computing*, 6(2), 52. <https://doi.org/10.3390/bdcc6020052>
- Stella, M., & Zaytseva, A. (2020). Forma mentis networks map how nursing and engineering students enhance their mindsets about innovation and health during professional growth. *PeerJ Computer Science*, 6, e255. <https://doi.org/10.7717/peerj-cs.255>
- Vlasceanu, M., & Amodio, D. M. (2022). Propagation of societal gender inequality by internet search algorithms. *Proceedings of the National Academy of Sciences*, 119(29). <https://doi.org/10.1073/pnas.2204529119>

ENGINEERING

Creating a Graphene Based Fluorescent Fingerprint Powder

Eamon McGarty

ABSTRACT:

Fingerprints are left everywhere, anything touched will leave oils from the fingers and could be traced back because they are unique, but how could they improve to better and more easily identify someone? When fingerprints aren't taken properly or identified, it causes a problem that contributes to the loss of evidence or reduction in solving crimes. This project is looking at manners to improve fingerprint powders to enhance visibility and sensitivity through a novel method, using new materials to create a different solution, such as mixing graphene powder (a different form of carbon made up of a single layer of atoms) and varying amounts of ethanol (independent variable) to see if the fingerprint powder could improve by being able to identify more minutiae. The amount of ethanol added to the solution varied by quantity at 50mL, 75mL, and 100mL and then measured the number of minutiae identified (dependent variable) in a fingerprint through a software that detects minutiae. The hypothesis was that the integration of graphene powder would significantly enhance the sensitivity (ability to detect prints) and selectivity (ability to distinguish between different prints) of latent fingerprint detection when compared to traditional fingerprint powders. The experimental results didn't support the hypothesis because the P-Values were calculated, however not included in the paper. This is because an average human fingerprint has around 25 to 80 minutiae, and with the novel fingerprint powders created in this study having over 600 minutiae, it's not possible for the data to be accurate.

INTRODUCTION:

The purpose of this project is to create a more easily applicable method for enhancing fingerprint detection, potentially revolutionizing the precision and accuracy of fingerprinting to help at crime scenes and forensic analysis.

Fingerprint analysis has long been the basis of criminal investigations, aiding in the identification and conviction of offenders for over a century. The uniqueness and permanence of fingerprints, combined with their commonality at crime scenes, make them a reliable biometric identifier. There are many ways to collect and identify fingerprints such as magnetic carbon powders, fluorescent powders, microsil, and cyanoacrylate. Using these techniques, it is possible

identify one of the eight fingerprint classifications: Plain Arches, Tented Arches, Ulnar Loops, Radial Loops, Plain Whorls, Central Pocket Whorls, Double Loop Whorls, and Accidental Whorls. A study done by Shrestha & Malla shows that loops are the most common type, with 52.71% of the 196 individuals in the representative sample having them, the second being the whorl, with 27.38% of the representative sample having them, and the rarest, the arch, with only 15.28% of the representative sample having them (Shrestha & Malla, 2019). It is also possible to identify minutiae, which are classified into one of the eight patterns: bifurcations, deltas, ridge endings, enclosures, short ridges, islands, ridge crossings, and cores. Using a combination of both fingerprint classifications and minutiae patterns, they become a

crucial piece of evidence in criminal investigations since they are unique to each person, serving as a reliable means of individual identification.

However, the reliability of fingerprint analysis has been impaired by the inherent error rates influenced by human factors and the varying quality of fingerprints. A study has shown that fingerprint examiners have a “false positive error rate of 0.1% and a false negative rate of 7.5%.”, influenced from human error and more importantly, the quality of the fingerprint (Gong, 1). These statistics emphasize the critical need for precise processing techniques in this field.

Advancements in materials science have fueled the exploration of innovative materials for fingerprint analysis. Graphene, a two-dimensional carbon allotrope, has attracted attention not only for its electrical, magnetic, and thermal properties but also for its remarkable optical attributes (Mbayachi et al., 2021). The study specifically highlighted graphene's optical uniqueness, revealing its potential application in forensic science. The thinness of the two-dimensional carbon allotrope, which is the unique way the carbon atoms are arranged, allows light to pass through easier, making it a promising candidate for integration into fluorescent fingerprint powder, promising substantial improvements in contrast and sensitivity for latent fingerprint detection.

In consideration of these historical and technological advancements, this research seeks to address the contemporary challenges in fingerprint analysis by exploring graphene's unique optical attributes. Prior research has been done on other materials such as graphite, however this project seeks to use new and different materials than those

already experimented. By discovering graphene's potential in fingerprint powder, the research aims to improve the precision and reliability in forensic science and criminal investigations by improving the number of minutiae detected from a fingerprint powder. It is hypothesized that the integration of graphene into fluorescent fingerprint powder will significantly enhance the accuracy and reliability of fingerprint examination within forensic applications, when compared to traditional fluorescent fingerprint powders.

METHODOLOGY:

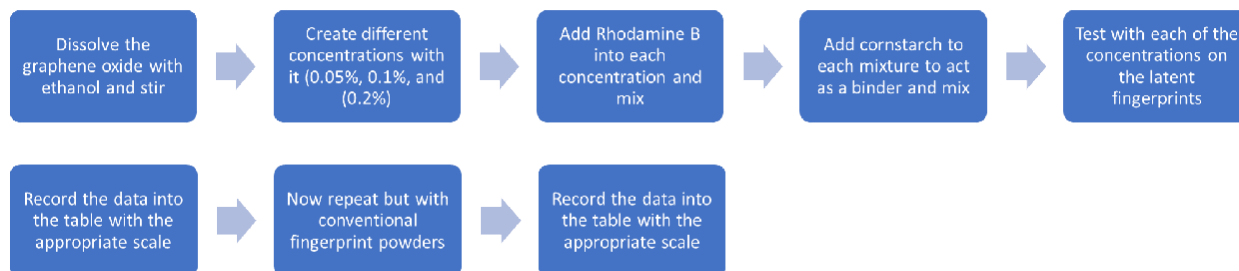
The materials required for the experiment include: 10 gram of graphene oxide powder (Brand: XFnANO), 1 gram of fluorescent dye (Rhodamine B, Brand: HiMedia), 1 gram of adhesive (Cornstarch, Brand: 365 by Whole Foods Market), 300 mL of solvent (Ethanol 70%, Brand: Amazon Basics), fingerprint samples from Eamon McGarty's right index finger, conventional fluorescent fingerprint powders (Brand: Sirchie, Orange Magnetic Fingerprint Powder, Fingerprint Minutiae Viewer Software (FpMV) (Brand: NIST), a phone (Google Pixel 7a), photo editing software (Adobe Photoshop), and four pages of 8.5 inches by 11 inches blank printer paper (Brand: Up & Up).

The design protocol begins by gathering all the necessary materials. Afterwards, 1 gram of graphene oxide suspension is dissolved into three different solutions of ethanol, 50 mL, 75 mL, and 100 mL of ethanol solvent, ensuring even dispersion through stirring in a beaker until all the materials are fully dissolved together. Then, fingerprint powders will be created with the varying amount of ethanol added. In each batch, 0.2 grams of fluorescent dye are dissolved into the graphene oxide suspension, or

otherwise known as a heterogeneous mixture. Thorough stirring is performed to achieve an even distribution of the dye and powder. Additionally, 0.2 grams of cornstarch are added to each batch as a

binder to enhance adherence, followed by another round of thorough stirring. The final step is to allow the solvent to evaporate overnight in the chemical hood.

Model 1.1:



Flow Chart of Developing and Testing Graphene Fluorescent Dye

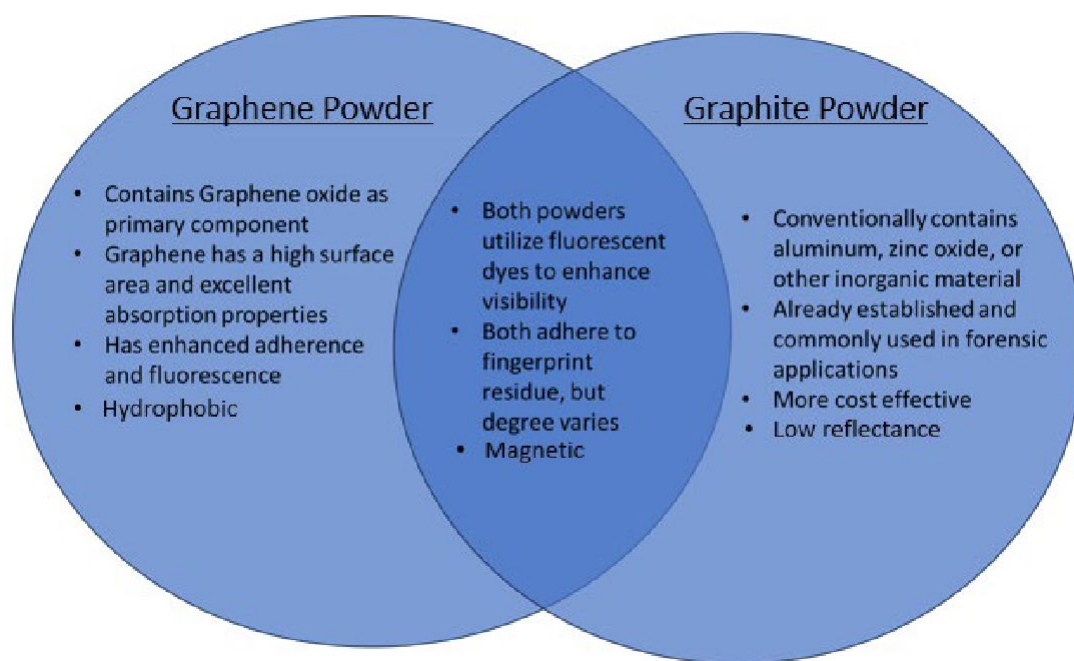
A simplified and more easily understandable flow chart of the process of developing and testing the graphene fluorescent dye.

The experimental procedure utilizes amount of ethanol in the fingerprint powder as the independent variable, with the amount of ethanol of 50 mL, 75 mL, and 100mL. The dependent variables include the number of minutiae, which was evaluated through the FpMV software by examining detectable minutiae with values above 0.15 quality, set though the software.

The experimental steps involve the preparation of the Up & Up blank printer paper with Eamon McGarty’s right index fingerprints. After, 1 gram of graphene-based samples (50 mL, 75 mL, and 100 mL) and 1 gram of conventional powders are applied to each fingerprint location. Ten trials of

fingerprints are taken for each powder type, all on 4 separate sheets of the printer paper. Photos of the fingerprints are taken using the Google Pixel 7a, and these photos are converted to grayscale using the adobe photoshop and saving the image as a JPEG. The grayscale photos are then inserted into the FpMV software, setting the quality threshold for minutiae to be greater than 0.15 in order to make sure the detected minutiae are visible enough to be able to identify. Finally, the number of minutiae is recorded, and are documented in tables 1.1, 1.2, 1.3, and 1.4, corresponding to the respective fingerprint powder.

Model 1.2:



Venn Diagram Comparing Graphene Powder Vs. Graphite Powder

A Venn diagram comparing and contrasting Graphene Based Fingerprint Powder and Conventional Based Fingerprint Powder.

Data Table 1.1 Graphite Based Fluorescent Powder (Control)

<i>Fingerprint Trial</i>		1	2	3	4	5	6	7	8	9	10
<i>Number of Minutiae Detected</i>		41	43	35	31	42	40	33	30	32	35

Data Table 1.2 Graphene Based Powder (50 mL Ethanol) (Experimental 1)

<i>Fingerprint Trial</i>	1	2	3	4	5	6	7	8	9	10
<i>Number of Minutiae Detected</i>	684	704	653	642	788	757	729	675	691	736

Data Table 1.3 Graphene Based Powder (75 mL Ethanol) (Experimental 2)

<i>Fingerprint Trial</i>	1	2	3	4	5	6	7	8	9	10
<i>Number of Minutiae Detected</i>	837	813	856	872	794	847	785	819	884	825

Data Table 1.4 Graphene Based Powder (100 mL Ethanol) (Experimental 3)

<i>Fingerprint Trial</i>	1	2	3	4	5	6	7	8	9	10
<i>Number of Minutiae Detected</i>	852	799	843	833	871	863	859	892	828	907

CONCLUSION:

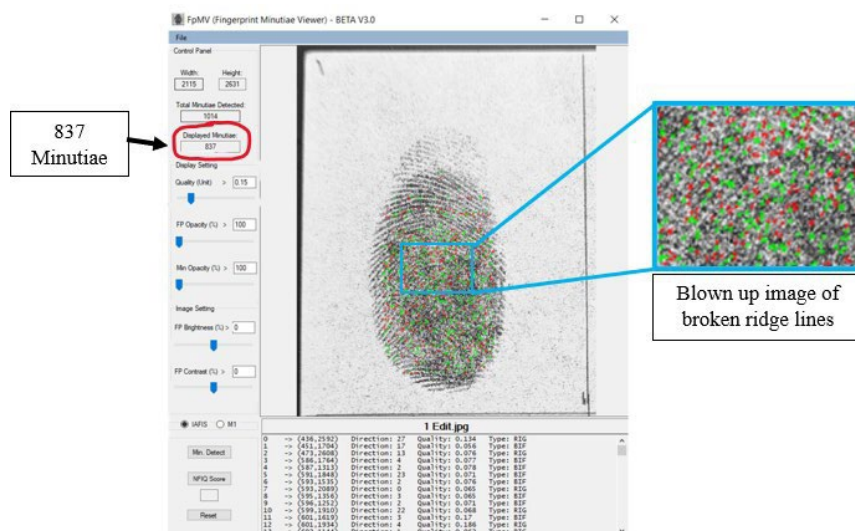


Figure 1: The displayed results of the number of minutiae found on the first trial with the 75 mL ethanol solution inside the FpMV software

Using excels features, calculate the mean and standard deviation of each of the trials of fingerprint powders. After that, by comparing each of the created fingerprint powders which varied in ethanol concentrations of 50 mL, 75 mL, and 100 mL of ethanol to the traditionally used fingerprint powder, calculate the P-Values. The P-Values are

calculated by comparing the observed data, which in this case, is the traditionally used fingerprint powder, to the expected data, which are the formulated fingerprint powders with varying amounts of ethanol, however, are not included in this paper because the number of minutiae detected were not humanly possible. This is because an average human

fingerprint has around 25 to 80 minutiae (J. R. et al., 2009), and with the novel fingerprint powders created in this study having over 600 minutiae, it is not possible for the data to be accurate. This could potentially be caused by the powder not being fine enough, breaking up the ridge lines and leaving gaps that the FpMV software detects as minutiae, leading to such a high number. As you can see in Figure 1 referenced above, there are big clusters of false positive minutiae scattered across the ridge lines, showing how gaps caused by the powder on the ridge lines could be what caused this high number of minutiae to be detected by the software.

Due to the newly created fingerprint powders not being accurate enough, they are not yet applicable as fingerprint powder, and some factors that could have contributed, including the amount of ethanol used. By calculating the mean of each of the novel fingerprint powders, the results were 705.9 minutiae for 50 mL ethanol, 833.2 minutiae for 75 mL ethanol, and 854.7 minutiae for 100 mL ethanol. The standard deviation was also calculated for each of the powders, with 46.35 for 50 mL ethanol, 32.23 for 75 mL ethanol, and 31.56 for 100 mL ethanol. With these results, it is shown that the solution created with the least amount of ethanol, 50 mL, had

the least number of average minutiae. This can lead to a future study done with even less ethanol, such as 25 mL or even down to 5 mL of ethanol to find if this could reduce the number of false positives found in the current fingerprint powders. As more trials are done, this lower amount of ethanol could also lead to a lower standard deviation, as increasing the number of trials in testing typically reduces the standard deviation. Another future step that could be taken is reducing the amount of graphene oxide powder added into the solution, as this could be the reason why the powder might not have been fine enough causing the gaps in the ridge lines. One more future step that could be taken with this project is by increasing the amount of cornstarch added to the solution, as this cornstarch was added for its adhesive properties. By increasing the amount of cornstarch from 0.2 grams to 0.5 grams, the increased cornstarch in the solution could contribute to the powder being able to better adhere to the surface with the fingerprint, leading to a greater fineness of the powder created and less minutiae being detected in the FpMV software. Therefore, more research needs to be conducted in order to find ways to reduce the number of detected minutiae.

BIBLIOGRAPHY:

- Bhati, K., Tripathy, D. B., Dixit, A. K., Kumaravel, V., Sabir, J. S. M., Rather, I. A., & Shukla, S. (2023). Waste Biomass Originated Biocompatible Fluorescent Graphene Nano-Sheets for Latent Fingerprints Detection in Versatile Surfaces. *Catalysts*, 13(7), 1077. <https://doi.org/10.3390/catal13071077>
- Dong, X.-Y., Niu, X.-Q., Zhang, Z.-Y., Wei, J.-S., & Xiong, H.-M. (2020). Red Fluorescent Carbon Dot Powder for Accurate Latent Fingerprint Identification using an Artificial Intelligence Program. *ACS Applied Materials & Interfaces*. <https://doi.org/10.1021/acsami.0c01972>
- Mbayachi, V. B., Ndayiragije, E., Sammani, T., Taj, S., Mbuta, E. R., & Khan, A. Ullah. (2021). Graphene synthesis, characterization and its applications: A review. *Results in Chemistry*, 3(3), 100163. <https://doi.org/10.1016/j.rechem.2021.100163>
- Ravi, J., Raja, K., & Venugopal, R. (2009). FINGERPRINT RECOGNITION USING MINUTIA SCORE MATCHING. *International Journal of Engineering Science and Technology*, 1(2), 35–42. <https://arxiv.org/ftp/arxiv/papers/1001/1001.4186.pdf>
- Shrestha, I., & Malla, B. K. (2019). Study of Fingerprint Patterns in Population of a Community. *Journal of Nepal Medical Association*, 57(219). <https://doi.org/10.31729/jnma.4621>
- Wang, M., Li, M., Yu, A., Zhu, Y., Yang, M., & Mao, C. (2017). Fluorescent Nanomaterials for the Development of Latent Fingerprints in Forensic Sciences. *Advanced Functional Materials*, 27(14), 1606243. <https://doi.org/10.1002/adfm.201606243>
- Zhang, M., Zhu, Y., Yu, X., Liu, S., Wang, M., Wei, Q., Hu, X., Tang, Q., Zhao, Y., & Zhang, X. (2013). Application of Electrodepositing Graphene Nanosheets for Latent Fingerprint Enhancement. *Electroanalysis*, 26(1), 209–215. <https://doi.org/10.1002/elan.201300412>

Development of Application for Early Detection of Oral Issues Based on Intraoral pH.

Nikhil S. Kadam

INTRODUCTION:

This project seeks to prove the correlation between salivary pH and intraoral dental issues, with the additional development of a prototype application which scans pH strips to analyze and identify salivary pH, maintaining a comprehensive history of pH changes over time. By providing users with insights to enhance their dietary choices, the application promotes the restoration of the user's salivary pH to a healthy range between 6-7.5, thereby reducing the likelihood of tooth decay. Oral health is a vital aspect of an individual's overall well-being, and the increased prevalence of tooth decay and dental caries (cavities) among adults establishes the need for innovative approaches to dental care. Dental caries, defined as a chronic infectious disease resulting from tooth-adherent bacteria, present a significant challenge as these bacteria metabolize sugars to produce acid which demineralizes tooth structure over time (Rathee, June 2023). According to the National Institution of Health Center, approximately 90% of adults aged 20 to 64 encounter issues related to tooth decay and caries. This issue worsens with age, with a staggering 70.1% of adults aged 65 and older experiencing gum infections.

When patients visit their dental practices, the doctors conduct thorough examinations for visual evidence of caries or tooth decay. However, the public lacks a reliable method to predict their

oral cavity's susceptibility to these issues, without assistance from the oral professionals. Though the dental industry has advanced in terms of declines in the prevalence and severity of caries, specifically in child populations, studies have shown a lack of change in “intraoral levels of cariogenic organisms or the annual consumption of sugars” (Burt, June 1998). Further research reveals that specific types of these cariogenic micro-bacteria play a crucial role in tooth decay, thriving at particular pH levels in the oral cavity. Notable micro-bacteria, including *Pophyromonas Gingivalis*, *Fusobacterium Nucleatum*, and *Streptococcus Mutans*, flourish at the critical pH range between 5.5 and 5.7 (Timpel, Sept 2023). This pH range has been scientifically proven to directly influence caries growth.

Despite the salivary pH stabilizing back to its resting range of 6-7.5 after proper oral cavity cleaning, individuals face a common challenge – the lack of awareness regarding their salivary pH after daily brushing, a crucial indicator of oral cavity health. Research done at the Department of Periodontology at Gian Sagar Dental College shows that out of the 1000 patients analyzed, a startling 56.8% had a severe lack of knowledge regarding the relationship between oral health and systemic health (Kapoor). Additionally, individuals lack the knowledge to analyze the effect of their nutritional food intake on their intraoral pH, and how this ultimately influences their oral cavity health.

To address this gap, this project proposes the development of an innovative application which allows individuals to take control of their oral health. This application allows users to analyze a pH strip, providing real-time monitoring of their salivary pH. The data can be entered manually or scanned directly from a pH strip, as seen in Figure 1. As of today, apps which already scan a pH strip do not focus on dental health, are unable to maintain historical data, and are subscription based, aspects which this application, named the pH-Tracker, implements thoroughly. By collecting and analyzing historical data, the application offers personalized recommendations, such as options for healthy acidic or alkaline foods to stabilize the intraoral pH back to its resting state; this guidance provides awareness of solutions to minimize the risk of caries and tooth decay. To ensure the concept of the app would be useful, the project analyzes patient pH data to establish correlations between pH levels and the occurrence of oral issues. The hypothesis for this study was an individual's pH negatively correlates with the severity of their oral cavity issues. The collection of data to support this hypothesis and development of the application were two separate parts of this research study.



Figure 1: pH Strips from JustFitter

METHODOLOGY Part 1: Data Collection / Analysis

In the initial phase of the project, proving the relationship between pH and oral issues was the first priority. To complete this task, dental offices provided two pH values from each patient willing to participate in the study, for consistency, along with historical patient data for further analysis. The individual data samples included dental observations and notes which the dental office had gathered throughout the patient's time at the office, such as periodontal disease, number of cavities, percentage plaque level, gum sensitivity, and other systemic issues. The data was completely anonymous, using a patient ID for each of the 15 patients to ensure identity without personal information documentation. To select patients, however, exclusion criteria was used to ensure the pH values were not skewed. Patients with other types of systemic diseases, such as Parkinsons disease, Breast Cancer, or Bite Force Pressure cavities, not related to oral health were excluded, as these diseases may cause patient pH to fluctuate, without the direct cause of tooth decay. For example, an individual with Parkinson's disease would have trouble brushing due to the lack of control in hand movements, which fluctuates their pH but would not be due to cavities. The collection of this patient data enabled analysis of the hypothesis: the negative correlation between the severity of oral issues with intraoral pH.

The first oral issue analyzed was the relationship between periodontal disease and pH. Periodontal disease is defined as the result of infections and inflammations of the gums and bone that surround and support the teeth. Each patient

which visits a dental office receives specific treatment based on their history with periodontal disease. An individual with no past history of periodontal disease receives Prophy treatment and is a patient with normal oral health. An individual with a history of periodontal disease receives Perio treatment. The data was analyzed by developing a

separate Excel sheet and deleting unwanted columns, such as Patient ID, date of data collection, and number of cavities present. Remaining was the patient's pH value and their corresponding treatment. The data was then split into two categories, patients with an acidic pH (below 7), and patients with an alkaline pH (7 or above).

Acidic pH			Neutral / Alkaline pH		
pH value	Perio?	Prophy?	pH value	Perio?	Prophy?
5.5			7		
5.5			7		
5.5			7.25		
5.75			7.25		
6			7.5		
6.75			7.5		
6.75			7.5		
			8		

Figure 2: Intraoral pH value vs. Perio/Prophy Treatment

The second oral issue analyzed was the relation between cavities and pH. This was accomplished using the cavity/year metric, which is a patient's number of cavity restorations divided by the number of years their data has been collected in

the dental office's database. As done with the periodontal disease study, the Excel sheet removed unneeded variables, focusing on the cavity/year and pH of each patient, as seen in Figure 3. Next, using the Excel Regression Graph feature, a dot plot was graphed.

Acidic		Neutral/Alkaline	
pH Value	cavities / year	pH value	cavities / year
5.5	1.1	7	0.05
5.5	1.3	7	1
5.5	0.8	7.25	0.7
6.5	1.2	7.25	0.6
6	0.3	7.5	0.3
6.75	0.28	7.5	0.4
		7.5	0.4

Figure 3: Data Table containing cavities/year vs. pH

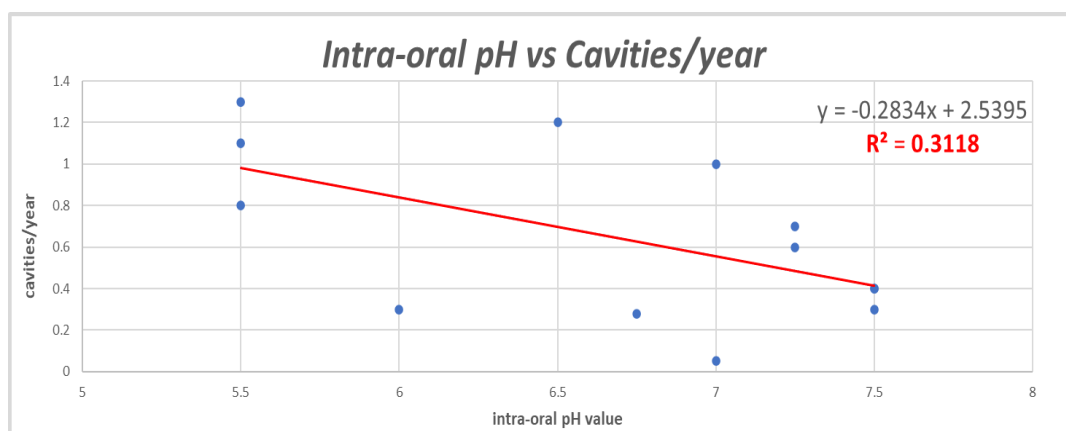


Figure 4: Intra-oral pH vs. Cavities/year

RESULTS AND CONCLUSION:

In the Periodontal Disease vs. pH analysis, 71% of the patients with an acidic pH were on Perio treatment while 100% of patients with an alkaline pH were on Prophy treatment. This supported the hypothesis because a lower, acidic pH corresponded with a higher susceptibility to periodontal disease. In the Cavities vs pH analysis, using the Regression Graph in Excel, a R value of -0.5583 was obtained, which is a negative correlation. As seen in Figure 4, the R^2 value of 0.3118 supports a strong negative correlation between cavities and intra-oral pH. To further support the correlation, a p-value calculated between pH and cavities. A p-value correlation coefficient is used to prove a positive or negative correlation between two variables. Using the Excel p-value function ($=T.TEST(array\ 1, array\ 2, tails, type)$), a p-value of 0.047 was obtained, which is statistically significant. This supports the hypothesis as well, as a negative correlation between pH and cavities was shown.

To summarize the results of both data analysis, periodontal treatment, and cavities, both showed a negative correlation between pH and intraoral issues. In other words, lower pH values

(acidic) had higher susceptibility for cavities and periodontal disease. From these findings, we can conclude the original hypothesis was supported.

Some limitations of this study include a small data sample, due to time constraints during data collection. Since 15 patients were used for the study, the reliability of the data is not as strong, because a larger sample provides more accurate results. However, as we can see from the line of best fit in Figure 4, a negative correlation is still shown, which should remain consistent if a larger sample set was obtained.

METHODOLOGY Part 2: Application Development

Now that the relationship between pH and oral issues is supported, the development of the application starts. As a brief overview, the application scans the pH strip to obtain the pH value, records the historical pH data of the user, offers improvement suggestions to the user on their oral habits, and ultimately boosts awareness on what an individual's pH indicates about their oral cavity health. One major benefit includes individuals taking ownership of their oral health by regularly

monitoring pH, thus preventing potential dental problems.

To develop the application, Visual Studio 2022 was downloaded for C# coding. Next, download the .net Maui sdk library, which stands for Multi-Platform App User Interface, which allows the implementation of the application onto multiple platforms, including iOS, Android, and Windows. The open source UI component library MudBlazor was also downloaded, working with .net Maui to generate multiple graphics onto your applications, such as layout grids, carousel swiping, line graphs, and tables. The Camera View and Colorful library were also downloaded to develop the pH scanner feature of the application. Finally, a smartphone was prepared with an iOS developer license for coding and testing purposes of the app.

First, for developing the main graphics for the application, the MudBlazor library was implemented. This library contained multiple

features which allowed certain visual aspects of the application to catch the attention of the users. The main features the application used were called the MudContainer and MudGrid features. The MudContainer feature was implemented to group multiple graphics onto the screen. For example, if the font of an entire MudContainer was changed, all the objects inside the container would change their text to that specific font. This allowed quick changes to be implemented easily into the application. The MudGrid feature was a basic grid display on the screen, which decided how many pixels each object took on the screen. For example, as seen in Figure 5, the first row's graphics took 100% of the screen, the second row's graphics took 50% of the screen, and the third row took 25% of the screen. By allowing each graphic's pixel percentage to remain constant, the application's graphics to be transferred between different software's, such as from phone to PC, as seen in Figure 6.



Figure 5: example of MudGrid which displays the percentage changes of each object on a graphic screen. For example, the top row takes 100% width of the screen, the second takes 50% width, and the third takes 25% width.

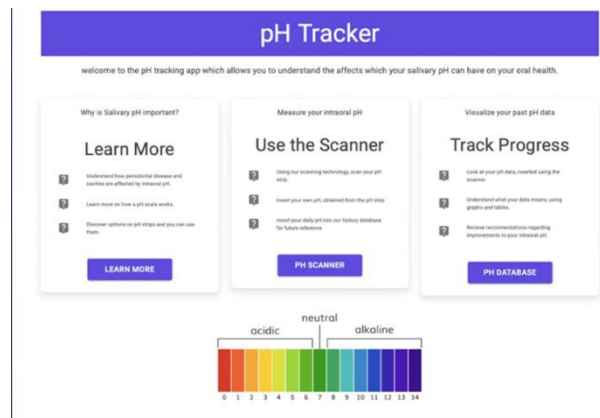


Figure 6: MudGrid example displaying a PC view, showing the consistency of each graphic object.

Next, the implementation of the pH scanner was the most challenging part of the project. A raw picture was first taken from the pH strip box, containing the key to the pH values, as seen in Figure 7. Using a Color Picker tool on Mac OS, each pH value's RGB values were obtained from the pH box key, and the values were stored in a list in C#, as seen in Figure 8. A challenge which was faced when developing the scanner, however, was a change in lighting in each room, which affected the color shown in the pH strip scan. To combat this issue, the library called Colorful was used, which contained a range of algorithms used to

compute a difference between color values. Specifically, it uses the color spectrum to determine the distance between two colors. The lower the distance, the closer the colors are on the spectrum. In the algorithm used for the application, the hundred-pixel squares on the center of each pH box were focused on, acting as a bar code scanner for the strips. After averaging each RGB value in that square, the Colorful library was used to determine the difference between the colors on the key. The match with the lowest difference was the value which the algorithm chose.



Figure 7: Raw picture of pH key used to determine RGB values of each pH value.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1			4.5	5	5.5	5.75	6	6.25	6.5	6.75	7	7.25	7.5	8	8.5	9
2		Red	229	221	219	217	220	222	216	185	166	130	99	58	45	12
3	Top Box	Green	93	93	109	114	126	138	143	128	118	102	89	63	61	35
4		Blue	0	0	0	0	0	0	0	0	0	0	0	0	17	44
5		Red	234	210	163	133	99	76	67	62	56	52	45	32	16	19
6	Bottom Box	Green	150	140	116	102	76	65	60	56	50	50	48	39	35	28
7		Blue	0	0	0	0	0	0	0	0	1	9	12	16	25	37

Figure 8: Excel sheet containing all the RGB values for each pH value from the key. Used in a C# list to compare each value to the pH strip value to determine the pH of the patient.

Some limitations for the development of the pH-Tracker Application included the use of only one pH strip brand. Since each pH strip brand is different in terms of the key they use to determine salivary pH, the application is unable to be compatible with multiple brands. However, for future research, the development of an option to pick which pH brand

you use will be implemented. Additionally, for future research, this project plans to work with dental offices to help them implement the application for their patients. This will allow the doctors to remotely monitor their patients, with the use of their salivary pH. This eliminates the need for constant checkups and appointments.

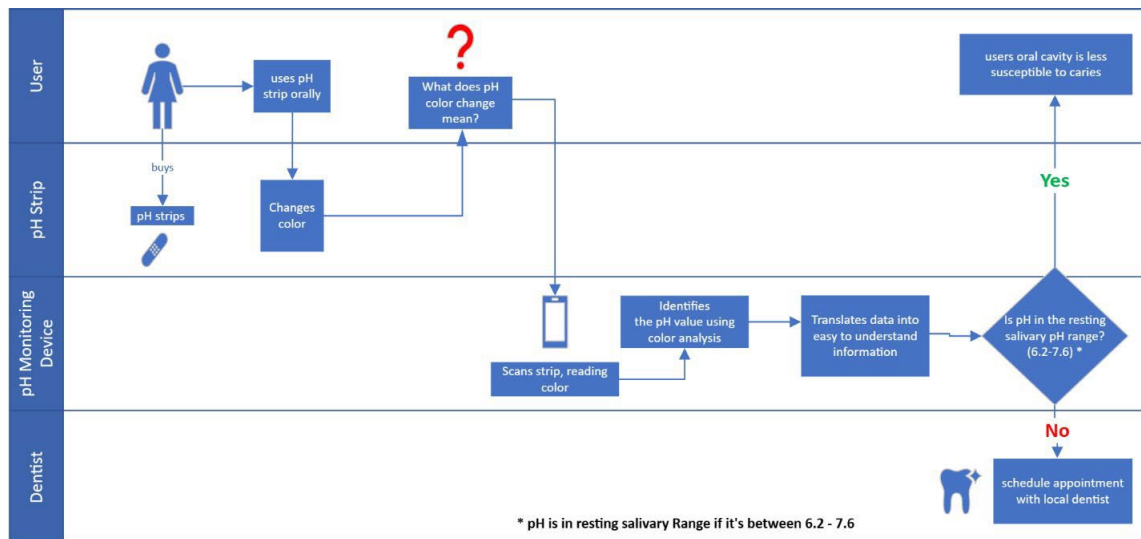


Figure 9: Flowchart explaining how the pH-Tracker application works from the user's perspective.

BIBLIOGRAPHY:

- Kapoor, Daljit. “Oral hygiene awareness and practice amongst patients visiting the Department of Perodontology at a Dental College and Hospital in North India”
National Library of Medicine, PubMed Central, 5 Apr. 2015,
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4184328/>
- Rathee, M., & Sapra, A. “Dental Caries” StatPearls, StatPearls Publishing, 21 Jun, 2023,
<https://pubmed.ncbi.nlm.nih.gov/31869163/#:~:text=Dental%20caries%20is%20a%20prevalent,over%20ti me%20demineralizes%20tooth%20structure.>
- Burt B. A. “Prevention policies in the light of the changed distribution of dental caries”, Actaodontologica Scandinavica, National Library of Medicine, 3 Jun, 1998, <https://pubmed.ncbi.nlm.nih.gov/9688230/>
- Farella, Mauro. “Simultaneous Wireless Assessment of Intra-Oral pH and Temperature.” Journal of Dentistry, Science Direct , 22 Mar. 2016, <https://doi.org/10.1016/j.jdent.2016.05.012./>
- Soni, Anuradha. “Smartphone Based Non-Invasive Salivary Glucose Biosensor.” Department of Biomedical Engineering, ELSEVIER, 20 Aug. 2017, <https://doi.org/10.1016/j.aca.2017.10.003./>
- Thompson, Clarence. “A Micro-Antimony Electrode Designed for Intraoral pH Measurements in Man and Small Experimental Animals.” Journal of Dental Research, American Association for Dental, Oral, and Craniofacial Research, Dec. 1954, <https://doi.org/10.1177/00220345540330061301./>
- Timpel, Julia. “Sensors for in Situ Monitoring of Oral and Dental Health Parameters in Saliva - Clinical Oral Investigations.” SpringerLink, Clinical Oral Investigations, 12 Sept. 2023,
<https://link.springer.com/article/10.1007/s00784-023-05206-9./>
- Xu, Zeqian. “Saliva-Coated Titanium Biosensor Detects Specific Bacterial Adhesion and Bactericide Caused Mass Loading upon Cell Death.” Biosensors and Bioelectronics, ELSEVIER, 23 Nov. 2018,
<https://doi.org/10.1016/j.bios.2019.01.035./>

Investigating the Effects of Biodegradable Plastic Breakdown on Arbuscular Mycorrhizal Fungi Colonization

Regan Hendrickson

ABSTRACT

As microplastics become more of a threat to natural environments and biodegradable plastics (BDPs) are increasingly being implemented, their effects on soil communities are largely unknown. This project seeks to investigate the effects that biodegradable plastic presence in soil (IV) has on sorghum plant growth (DV) and arbuscular mycorrhizal fungal (AMF) colonization (DV). It was hypothesized that there would be a statistically significant decrease in both sorghum growth and AMF colonization when plastic is present in soil. Three different fungal species were used in testing—*Rhizophagus irregularis*, *Rhizophagus intraradices*, and *Funneliformis mosseae*—as well as a control with no fungal presence. Half of the pots were treated with the addition of shredded biodegradable plastic mulch, while the other half was treated with a control of no plastic addition, totaling 8 different treatments. Roots were then harvested, dried, weighed, and digitally analyzed for fungal colonization. Through ANOVA analysis of sorghum growth, p-values greater than 0.05 were found and the null hypothesis (any change in sorghum growth is purely due to chance) was accepted. Fungal colonization percentage was calculated, and t-tests were performed to compare colonization in pots treated with and without plastics. These tests showed a statistical significance in fungal colonization based on plastic presence (R. intra WP and NP: $p = 8.068 \times 10^{-10}$; R. irreg WP and NP: $p = 1.547 \times 10^{-7}$). The hypothesis for this project was only partially supported – while BDP had no above ground impacts on sorghum growth, there was a drastic decrease in AMF colonization when plastic was present.

INTRODUCTION

In order to address decreasing soil health and the rapid implementation of biodegradable plastics as a replacement for single use plastics, this project seeks to directly investigate the effects that biodegradable plastics (BDPs) have on arbuscular mycorrhizal fungi (AMF) growth by measuring fungal colonization percentage of communities with and without BDP.

The symbiotic relationship that AMF form with their host plants support most of the plant life on Earth. In fact, mycorrhizal fungi form a mutualistic relationship with more than 80% of rooted plants by entering root tissue and building

fungal structures within the host plant. (Chalker-Scott 2009). In this relationship between the fungus and host plant, crops support the fungi's metabolic needs by supplying carbohydrates and lipids, and in exchange, fungi provide a myriad of benefits. These benefits include increased nutrient concentration, by way of more efficient nutrient absorption and a larger web of fungi than roots alone. (Zhang et. al., 2019). Fungi also have been shown to improve soil structure, stabilize carbon in soil aggregates, and improve nutrient exchange (Leifheit et. al., 2021). Since soils are intricately connected to Earth's aquatic and atmospheric systems, it is critical that further research is conducted to observe the effects

that common pollutants have on the health of beneficial microorganisms, such as mycorrhizal fungi, in soil communities.

One major threat to the health of soil communities globally is the accumulation of microplastics. The usage of single use plastics has only increased since they were first implemented in the late 1970s. Single use plastics like plastic bags and packaging are typically light, durable and have a lifespan of less than 6 months. These characteristics make them especially prone to becoming microplastics (MPs) when exposed to sunlight and weathering (Filiciotto 2021). These microscopic pieces are not easily removable from soil, water, and air and their accumulation threatens the world's biodiversity. They have been shown to cause air pollution and negatively affect human health, especially the respiratory system. Microplastics are especially harmful to soil health as they destroy the natural crumb structure and decrease soil pH.

One potential solution to this issue is the implementation of biodegradable plastics (BDPs), like polybutylene adipate terephthalate (PBAT) or polylactic acid (PLA). These plastics are meant to be broken down into organic molecules, which can then be assimilated by microorganisms and integrated into the soil. These benefits are why BDPs are widely supported as a sustainable alternative in multiple manufacturing sectors, including single-use items, packaging wrap, and agricultural plastic mulch films (Filiciotto 2021). However, many studies suggest that BDPs may break down into microplastics in unideal conditions, such as water, cold temperatures, and lack of sunlight (Qin et. al., 2021).

These findings are particularly relevant in understanding how biodegradable mulch films will break down in an agricultural setting – a review written by Aoyun Mo and their colleagues at East China Normal University suggests that “BDPs may generate lots of MPs and nanoplastics (NPs) and release chemicals” into the soil (2023). This research suggests that unlike traditional microplastics (which are considered chemically inert and only affect physical properties of the soil), biodegradable microplastics may alter the soil's physical and chemical compositions. This change in chemical composition has led to a theory that there is a negative correlation between the presence of BDP and AMF. A review written by Eva. F Leifheit references a study done by Fayuan Wang investigating the combined effects of microplastic and cadmium on plant growth and AMF growth: “biodegradable PLA increased the soil pH and such effects on soil pH can affect AMF community composition...” (Leifheit et. al. 2021).

There is little known about the effect of biodegradable plastic breakdown on mycorrhizal fungi colonization. In fact, there has been no experimentation on how the establishment of different species of AMF are impacted over time by the presence of BDPs and MPs in the soil. This project seeks to investigate how plant symbiotic relationships are impacted by the presence of these BDPs. A deeper understanding of these reactions could help drive better practices surrounding BDPs that focus on healthy soil, higher plant yields, and more efficient food production.

Based on this research, it can be

hypothesized that the presence of biodegradable plastic will significantly reduce the fungal

colonization of *Funnelformis mosseae*, *Rhizophagus irregularis*, and *Rhizophagus intraradices*.

MATERIALS AND METHODS:

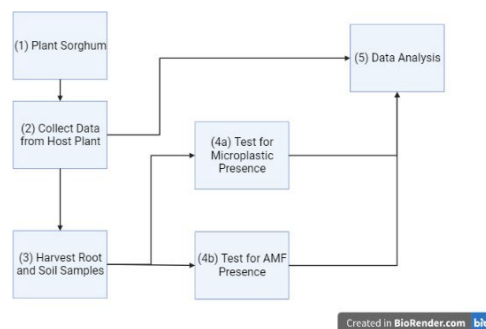


Figure 1: Experimental Design

First, the host sorghum seedlings were planted in eight different treatments, as detailed in *Table 1: Varying Treatments and Pot Numbers*, differing in species of AMF and presence of BDP. Then, physical data was taken from the plants—such as height and mass. Then root and soil samples were harvested from each pot. Data regarding microplastic presence was collected by measuring the mass of microplastics in comparison to mass of the soil sample. Data regarding AMF presence was collected by observing roots and fungi underneath a microscope. Finally, the data collected was analyzed via the procedure described in the Data Analysis section of this research paper. *Graphic created in BioRender.com by Regan Hendrickson.*

AMF and BDP

In this study, three species of arbuscular mycorrhizal fungi were used: *Rhizophagus irregularis*, *Rhizophagus intraradices*, and

Funnelformis mosseae. The biodegradable plastic used in this study is Oxo-Degradable Black Plastic Mulch 0.5 mil, a biodegradable mulch made by Grower's Solution.

Planting

Sorghum seeds were surface sterilized with 70% ethanol and household bleach and left to germinate in nutrient rich agar plates for 48 hours. Autoclaved agricultural soil, Sunshine Mix (a growing mixture that provides nutrient and water retention, among other benefits) and sand were mixed in a 1:1:1 ratio. 2 L of this mixture were measured out into thirty-two 1-gallon pots. Each pot was labeled twice – once on a planting label stick and the other on label tape on the side of the pot. The label included information on fungi species, presence of plastic, and assigned treatment letter, as described in *Table 1: Varying Treatments and Pot Numbers*.



Figure 2: Labeled Pot with Planting Stick and Label Tap
Each pot was labeled twice with the with the AMF present, presence of plastic, and treatment identifier. *Graphic created in BioRender.com by Regan Hendrickson*

Table 1: Varying Treatments and Pot Numbers

AMF Species	Presence of BDP	Treatment Identifier	Pot #
F. Mosseae	1.4 g	A	1-4
	None	B	5-8
R. irregularis	1.4 g	C	9-12
	None	D	13-16
R. intraradices	1.4 g	E	17-20
	None	F	21-24
N/A – Controls	1.4 g	G	25-28
	None	H	29-32

The biodegradable plastic mulch was shredded into random sized pieces and 1.4 g of plastic was measured and mixed into the top 1-inch of pots labeled with treatment A, C, E, and G.

6 grams of *F. Mosseae* was measured out and mixed into the top 1-inch of pots labeled with treatment A and B. The same quantity of *R. irregularis* and *R. intraradices* was added to pots labeled with treatment C/D and E/F respectively.

4-5 sorghum seeds of even emergence were planted in each pot. After 1 week of growth, if there were more than 3 successful seedlings, pots were randomly thinned to only have 3 plants.

Plants were kept in one of the experimental greenhouses at the University of Washington. Relative humidity was maintained within 55-66% and temperature was maintained within 68-76° F. To control variation in sunlight, air flow, and other

extraneous factors, pots were arranged in a randomized block design. This arrangement included one replicate of each treatment in each row and column. Soil moisture was maintained at field capacity (the maximum amount of water that soil can hold after excess water has drained away) for the whole growing period.

DATA COLLECTION

On day 32, the host plant height was measured and recorded in Table 2: Height of Sorghum Plants on Day 32. All plant matter was harvested from soil, and roots and above ground plant material (shoots) were separated. The plant biomass was kept in a drying oven for 12 hours in order to remove all moisture. Dry roots and shoots were individually weighed and recorded in Table 3: Dried Sorghum Plant Matter.

Table 2: Height of Sorghum Plants on Day 32

		Treatment							
Height (cm)	Pot	A	B	C	D	E	F	G	H
	1	41.5	30.7	39.9	43.3	77.4	63.1	54.7	44.5
		46.9	49.4	72.5	59	56.9	62.9	40.1	33.4
		35.7	44.1	49	78.3	39.7	52.3	50	51.4
	2	70.8	77	70.3	19	64.7	71.7	52.6	78.9
		53.3	42.3	45.6	57	57.9	60.6	74.7	48.3
		43.1	50.7	69.1	15.2	43.1	44.5	46.5	47.9
	3	38.7	43.3	30.5	66.4	73.5	60.6	56.7	86.2
		62.7	55.5	66.5	39.1	31.5	63.2	63.1	30.4
		32.1	34	37.1	54.5	48.2	56	44	17.9
	4	13.5	44.6	45.3	38.8	47.5	49.9	52	48
		22.7	38.5	36.7	53.8	31.9	57.5	59.5	62.6
		63.2	59	61.5	52.7	25.3	64.4	60.1	21.9

Table 3: Dried Sorghum Plant Matter

	Treatment							
	A	B	C	D	E	F	G	H
Foliage Weight (g)	0.51	0.71	0.87	1	0.97	1.39	1.54	0.56
	1.96	0.78	0.51	0.69	1.83	1.05	0.83	1.35
	1.96	0.45	0.49	2.08	1.16	1.61	1.54	1.63
	0.75	1.93	0.55	0.69	0.94	2	1.64	2.44
Root Weight (g)	0.1	0.1	0.15	0.03	0.13	0.26	0.26	0.12
	0.11	0.07	0.37	0.12	0.36	0.08	0.18	0.14
	0.1	0.06	0.44	0.4	0.19	0.15	0.48	0.13
	0.06	0.26	0.06	0.48	0.14	0.04	0.28	0.22
Combined Weight (g)	0.61	0.81	1.02	1.03	1.1	1.65	1.8	0.68
	2.07	0.85	0.88	0.81	2.19	1.13	1.01	1.49
	2.06	0.51	0.93	2.48	1.35	1.76	2.02	1.76
	0.81	2.19	0.61	1.17	1.08	2.04	1.92	2.66

Then, pigment was cleared from roots using the procedure shown in *Figure 3: Clearing and Staining Roots*. Fine roots (<2mm diameter) were removed from the plant, placed

in a 2mm sieve and rinsed in MilliQ water to remove dirt and particulate matter. Tissue cassettes were labeled with the treatment identifier and pot number. 0.15 grams of roots

were measured and placed in the corresponding tissue cassette – samples were kept in separate containers, and this process was repeated for

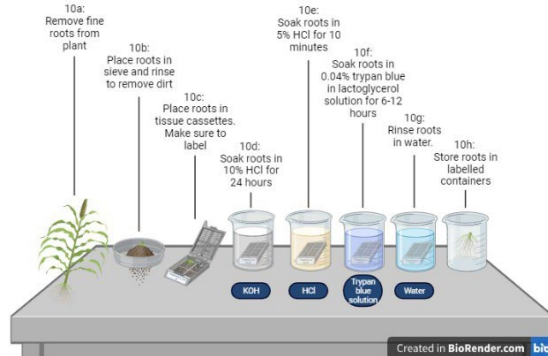


Figure 3: Clearing and Staining Roots

The clearing and staining roots procedure will remove all the natural pigment from the roots and dye the fungal cells to be identifiable under a microscope. Created in BioRender.com by Regan Hendrickson.

To obtain an objective measure of fungal colonization, an updated version of the gridline intersect method, as described in McGonigle et. al. 1990 was used. Fine root segments were then placed under a stereomicroscope and an image was taken. The CZI file was imported into Fiji, and image processing package of ImageJ. Then, using the line and measure tool, a length of 1500 micrometers was marked and a grid with at least 100 intersections on the root was overlaid.

each pot. Roots were then soaked in 10% KOH for 24 hours, then 5% HCl for 10 minutes, and 0.04% trypan blue in lactoglycerol solution for 6-12 hours. Finally, roots were rinsed in MilliQ water and stored in labeled containers.

This CZI file was then exported to a JPEG and opened with the Paint application. Each intersection was analyzed to have intraradical hyphae, arbuscules, vesicle/spore, or nothing. The number of each category was then tallied and the fungal colonization percentage was calculated, as shown in Table 4: Arbuscular Mycorrhizal Fungal Colonization by Treatment.

These observations only occurred for *R. irregularis* and *R. intraradicis* since no colonization was found in the *F. mossae* or control pots. Three images were analyzed per pot, totaling to 12 images per treatment.

Table 4: Arbuscular Mycorrhizal Fungal Colonization by Treatment

	Treatment			
	C	D	E	F
Fungal Colonization Percentage	R. irregularis		R. intraradicales	
	WP	NP	WP	NP
	0	0.7	0.542857	0.778761
	0	0.613861	0.354167	0.721154
	0	0.58	0	0.712871
	0	0.14	0	0.580357
	0	0.413333	0	0.56
	0	0.503876	0	0.672131
	0	0.616541	0	0.660377
	0	0.476821	0	0.631902
	0	0.651163	0	0.651163
	0	0.51773	0	0.529703
	0	0.72973	0	0.592308
	0	0.454545	0	0.777778

RESULTS AND CONCLUSION

Data analysis was conducted in Microsoft Excel, Google Sheets, and RStudio. Data was first imported and saved in an Excel document. Complete data analysis can be found at this link: <https://bit.ly/ReganWSSEFdata>.

Plant Biomass

Using the mean, median, and mode functions in Excel, values were derived for each treatment, as shown in *Table 5: Sorghum Growth Measures of Central Tendency*. The standard deviation and standard error were also calculated and stored in

Excel. Then, the data was imported into Google Sheets, where an ANOVA test was conducted using the XLMiner Analysis ToolPak add-in. Through ANOVA testing, there was not a statistically significant ($p > 0.05$) difference on the plant matter (above/below grown biomass and plant height) between groups treated with and without plastic, meaning that the null hypothesis was supported, and the alternative hypothesis was rejected.

Table 5: Sorghum Growth Measures of Central Tendency

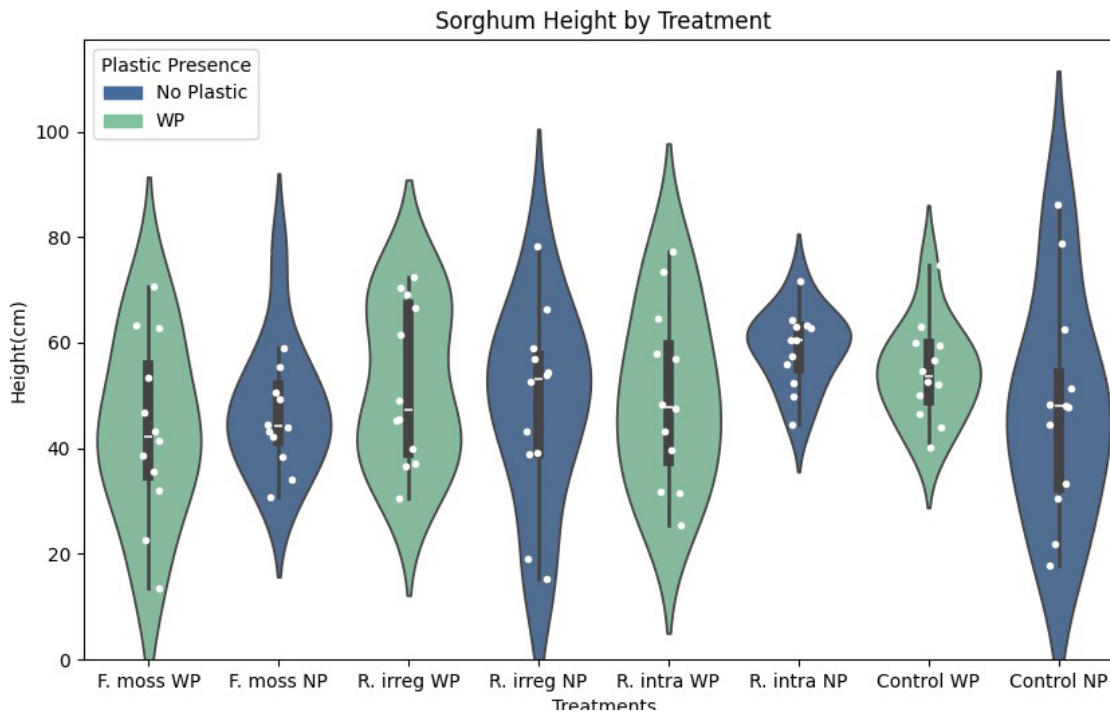
	A	B	C	D	E	F	G	H
Height on Day 32								
Mean	43.683	47.425	52.000	48.092	49.800	58.892	54.500	47.617
Median	42.300	44.350	47.300	53.250	47.850	60.600	53.650	47.950
Mode	#N/A	#N/A	#N/A	#N/A	#N/A	60.6	#N/A	#N/A
Shoot Dry Weight								
Mean	1.295	0.9675	0.605	1.115	1.225	1.5125	1.3875	1.495
Median	1.355	0.745	0.530	0.845	1.065	1.500	1.540	1.490
Mode	1.960	#N/A	#N/A	0.690	#N/A	#N/A	1.54	#N/A
Root Dry Weight								
Mean	0.093	0.1225	0.255	0.258	0.205	0.1325	0.300	0.153
Median	0.100	0.085	0.260	0.260	0.165	0.115	0.270	0.135
Mode	0.100	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Combined Dry Weight								
Mean	0.610	0.810	1.020	1.030	1.100	1.650	1.800	0.680
Median	2.070	0.850	0.880	0.810	2.190	1.130	1.010	1.490
Mode	2.060	0.510	0.930	2.480	1.350	1.760	2.020	1.760

Then, Mann Whitney U Tests were conducted in Excel to investigate the effect that plastic had on each fungi species. Like a T-Test, The Mann Whitney U Test is used to compare the difference between two groups. However, the T-Test assumes that the two groups are normally distributed and have similar variance, so the Mann Whitney U Test was used to account for the data being skewed and have different variances. When comparing the control group (no fungal presence) with plastic (WP) to the control with no plastic (NP), there was a

statistically significant difference ($p=0.0433$), with the WP trending higher than NP. This directly contradicted the hypothesis that there would be negative effects on plant growth with plastic present in the soil since root growth was greater when there was plastic.

However, this was the only comparison that showed a direct difference. For all other Mann Whitney U Tests conducted $p > 0.05$. Violin plots to represent this data were created in Python using the Seaborn and Matplotlib libraries.

Figure 4: Sorghum Height by Treatment



The height of the sorghum plants was taken on Day 32 and plotted on a violin plot in Python to show general trends in height. As the graph shows, there is no observable significant difference between any of the groups. This observation is supported through ANOVA analysis and Mann Whitney U Tests.

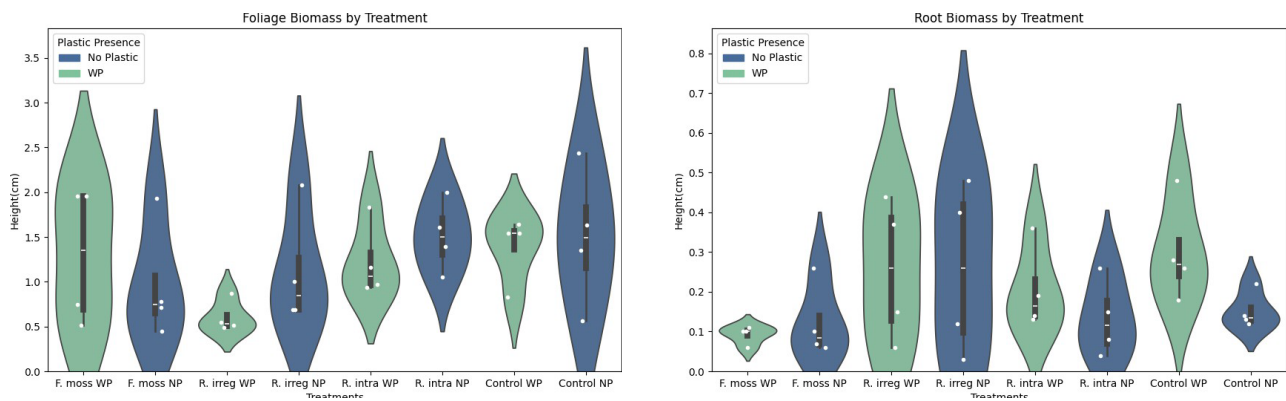


Figure 5: Foliage (a) and Root (b) Biomass by Treatment

After plants were harvested from the soil, the above ground plant matter (foliage/shoot) was separated from the below ground plant matter (root) and plotted on a violin plot in Python.

Through ANOVA analysis and Mann Whitney U Tests, there was no statistically significant difference between the groups. However, when visually comparing treatments with and without plastic in *Figure 5: Foliage and Root Biomass by Treatment*, there is a slight trend of foliage biomass being higher when plastic wasn't present in the soil (especially for *R. irregularis* and *R. intraradices*). This observable trend cannot be seen in the root biomass – in fact, it may appear that the root biomass was higher with a plastic presence in

the soil. While the reason for this is unclear, it is currently being hypothesized that the biodegradable plastic is being broken down into organic carbons that are feeding the root growth. *Figure 6: Total Biomass by Treatment* looks very similar to *Figure 5.a: Foliage Biomass by Treatment*. This can be attributed to the fact that the above ground biomass makes up a much higher percentage of the total biomass, so the effects that the plastic had on root growth may be masked when drawing conclusions based on total biomass.

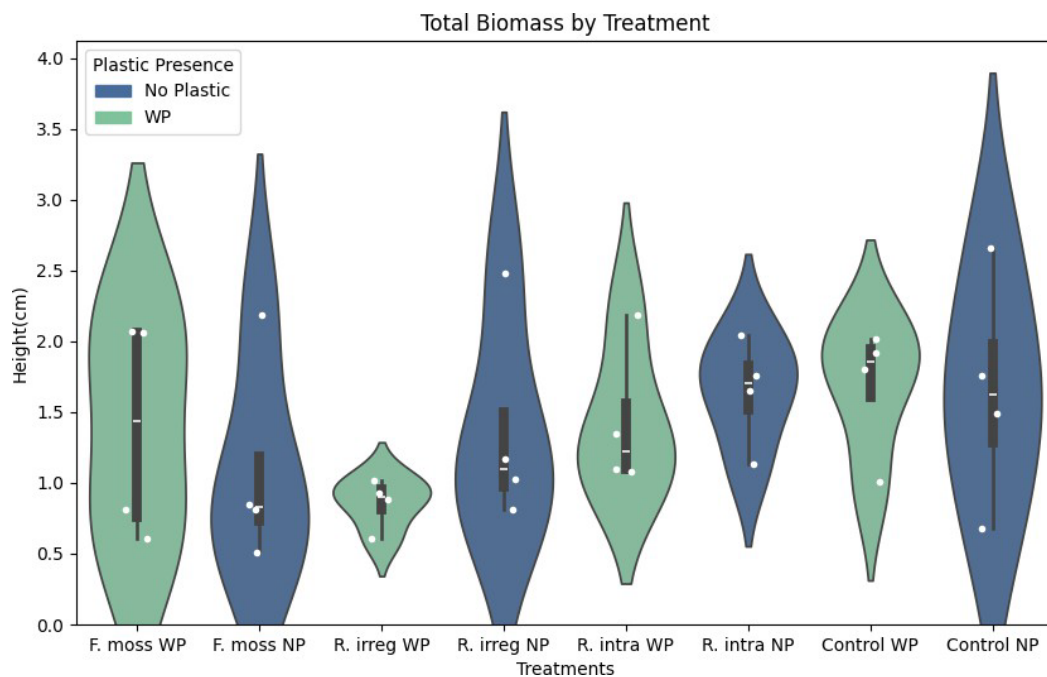


Figure 6: Total Biomass by Treatment

The above ground and below ground plant matters were added together in order to collect the total biomass. Using this data, a violin plot was made in Python.

Fungal Colonization

Data was imported into an Excel file, and using the mean, median, and mode functions, the measurements of central tendency were taken. Then, the data was imported into RStudio, and a T Test was conducted. When looking at *R. irregularis* and comparing WP to NP, $p = 1.547 \times 10^{-7}$, suggesting a statistically significant difference between the two groups. When comparing plastic presence in pots with

R. intraradices, $p = 8.068 \times 10^{-10}$. This supported the alternative hypothesis that when plastic is present in the soil, there would be adverse effects on fungal colonization.

In order to visualize the differences between the two groups, violin plots were made in RStudio. There was no observed colonization in the pots with *Funneliformis mossae*, so images were not analyzed with ImageJ and therefore, data regarding *F. mossae* was not analyzed.

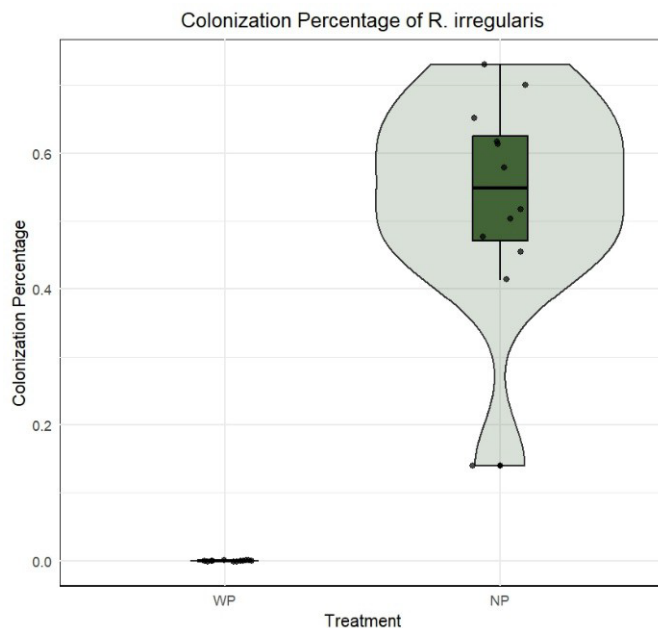


Figure 7: Colonization Percentage of *R. irregularis*

Data collected from the analysis of microscopic pictures was imported into R and violin plots were made. As shown on the left, there was no fungal colonization observed when plastic was present in the soil.

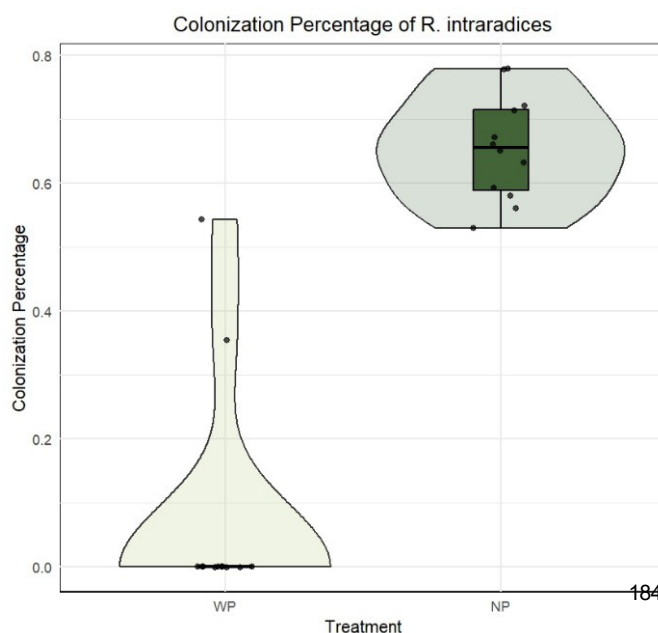


Figure 8: Colonization Percentage of *R. intraradices*

Similarly to [Figure 7](#), data was imported into R and violin plots were produced. As shown on the right, there were two instances in which colonization was observed in the treatment with plastic.

Unlike the pots with *R. irregularis*, there was some observed colonization when plastic was present. However, this only occurred in two of the twelve observed slides. While the reason for this is unclear, this could be attributed to a variation in how the plastic and fungal inoculum was mixed into the soil.

DISCUSSION

It was found that the presence of biodegradable plastic in the soil caused no reduction in host plant growth (as measured as biomass). However, there was an observed reduction in AMF colonization caused by plastic presence. This demonstrates the fungi's ability to protect the plant from abiotic stress factors. This is a critical finding because it shows that even if observations of above ground production show no obvious difference, there are detrimental impacts below the surface.

This study may have been impacted by the fact that the fact that plants were harvested before they completed their life cycle. An extended growing period would have given the fungi more time to colonize and sorghum more time to reach maturity. This change in experimental design could have two opposing effects. Had the sorghum reached full maturity, there may have been a more significant difference between the two groups. In contrast, a

more prominent fungal colonization may have boosted plant growth.

Another limitation to this study was the relatively small sample size. This size (only four pots per treatment) caused an increased probability of Type 2 error, as well as reduced power of statistical tests. This sample size particularly affected the data analysis of sorghum growth. In order to add more data points in regard to fungal colonization, more microscope images should be analyzed.

The next steps for this project are to investigate the soil content for microplastic and carbon contents. If there is a high microplastic presence in the soil, this may point to microplastic being the cause that colonization is reduced. If there is no microplastic presence in the soil, then additives being leached from the BDP may be causing the reduced colonization.

More research is also needed pertaining to the development of AMF in soil – especially how AMF develops when there is plastic present in soil. A longitudinal study regarding how soil communities change with pollutants present would provide some key insight into the hidden ecology of soil systems.

REFERENCES

- Chalker-Scott, L. (2009, season-04). Mycorrhizae: So, what the heck are they, anyway?
www.MasterGardenerOnline.com.
- Filiciotto, L., & Rothenberg, G. (2020). Biodegradable plastics: Standards, policies, and impacts.
ChemSusChem, 14(1), 56–72. <https://doi.org/10.1002/cssc.202002044>
- Grause, G., Kuniyasu, Y., Chien, M. F., & Inoue, C. (2022). Separation of microplastic from soil by centrifugation and its application to agricultural soil. *Chemosphere*, 288, 132654.
<https://doi.org/10.1016/j.chemosphere.2021.132654>
- Leifheit, E. F., Lehmann, A., & Rillig, M. C. (2021). Potential effects of microplastic on arbuscular mycorrhizal fungi. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.626709>
- McGonigle, T. P., Miller, M. H., Evans, D. G., Fairchild, G. L., & Swan, J. (1990). A new method which gives an objective measure of colonization of roots by vesicular—arbuscular mycorrhizal fungi. *New Phytologist*, 115(3), 495–501. <https://doi.org/10.1111/j.1469-8137.1990.tb00476.x>
- Mo, A., Zhang, Y., Gao, W., Jiang, J., & He, D. (2023). Environmental fate and impacts of biodegradable plastics in agricultural soil ecosystems. *Applied Soil Ecology*, 181, 104667.
<https://doi.org/10.1016/j.apsoil.2022.104667>
- Qin, M., Chen, C., Song, B., Shen, M., Cao, W., Yang, H., Zeng, G., & Gong, J. (2021). A review of biodegradable plastics to biodegradable microplastics: Another ecological threat to soil environments? *Journal of Cleaner Production*, 312, 127816.
<https://doi.org/10.1016/j.jclepro.2021.127816>
- Zhang, S., Lehmann, A., Zheng, W., You, Z., & Rillig, M. C. (2018). Arbuscular mycorrhizal fungi increase grain yields: a meta-analysis. *New Phytologist*, 222(1), 543–555.
<https://doi.org/10.1111/nph.15570>

O-Patch – Developing an Omega-3 Microneedle Patch for Women with Postpartum Depression

Harshita Raja

ABSTRACT

Postpartum depression (PPD) is a risk that anyone giving birth can face, which can affect mothers and families in negative ways socially, physically, and emotionally. A cheap and efficient microneedle patch with Omega-3 fatty acids will revolutionize the nutrition pathway for treating postpartum depression. The current problem is how to better treat postpartum depression knowing that a high dosage of 1.5 grams/day of fatty acids needs to be released, while being convenient and safe for recovering mothers. Firstly a 3-D printed mold was used in combination with Polyethylene Glycol and gelatin, which are heavily biocompatible, to create the actual patch and loaded with a payload solution of omega-3 fatty acids. This was then tested on a simulated skin model of gelatin for three variables for effectiveness: penetration depth of solution, release rate of the solution (RR) and total release amount (TAR). Gelatin and PEG are novel in that they act as a diffusive hydrogel, easily allowing the fatty acids to cross through the skin barrier with microneedles. The patches created had an average penetration depth of 1,319 micrometers, an average RR of 0.095 g/hour, and an average TAR of 2.25 g/day. This exceeds the engineering goals set for the investigation while having an average cost of \$15 per patch, making it affordable and efficient at diffusion. The p-values for each variable compared to industry standards show statistically significant data ($p = 0.004$ for penetration, $p = 0.018$ for TAR, & $p = 0.02$ for RR).

INTRODUCTION

This study aims to create and test an Omega-3 fatty acid microneedle patch for use on mothers with postpartum depression to aid recovery.

Postpartum depression (PPD) is a significant issue in perinatal women, with over 10% of them receiving a diagnosis. About 20% of all perinatal deaths come from postpartum depression, demonstrating the importance of addressing the problem. It has become an increasingly significant problem in urban communities, because mothers in this demographic are experiencing it more heavily than others. As countries and cities develop, the prevalence of PPD is likely to increase, which is daunting for future mothers. Though predicting postpartum depression is still beyond the reaches of science, finding viable

treatment options to improve the symptoms of the illness is very progressive. Polyunsaturated fatty acids (PUFAs), especially Omega-3 and Omega-6 polyunsaturated fatty acids, have been studied in terms of their benefits for women with postpartum depression. It has been shown to be indicative of whether a woman could potentially be diagnosed based on Omega-3 and Omega-6 serum PUFA levels after delivery (De Vriese et. al. 2003).

Following this, Omega-3 fatty acids and Omega-6 fatty acids have been found to help with depression. Though the exact mechanisms of how it occurs haven't been explained, a leading theory suggests that the fatty acids help with membrane fluidity in the brain, leading to an increase in efficiency of ion exchange, which is an important

part of action potential and synaptic firing. This can be extrapolated to PPD, since the mechanisms of depression have been seen to work similarly. Though there are several components of PPD, such as hormone levels and HPA-axis issues, the intake of these specific nutrients has been studied before using rodent models. In a study conducted by Camilla M.M da Rocha and Gilberto Kac, they found that a dietary ratio of the Omega-6 fatty acids and Omega-3 fatty acids should be as close to 1:1 as possible. It significantly decreased the prevalence of PPD after childbirth with a p-value of 0.04 (da Rocha and Kac, 2010). Studies like these have shown how important Omega-3 has been in reducing postpartum depression in mothers if taken during and after their pregnancy.

In a study conducted with female Long-Evans rats, who endured a separation PPD model, the intake of Omega-3 fatty acids through their diet was increased, and the expression of depression-like phenotypes were measured. The female rats, also known as dams, were separated repeatedly from their offspring, which caused stress and anxiety, leading to depressive-like phenotypes. This included lowered nursing behavior, increased time away from the nest, and lowered reaction times for several tests like the Forced Swim Test and Tail Suspension Test. After an integration of Omega-3 supplements in the diets of the rats over a period of 12 weeks, researchers saw a significant increase in the reaction times of the dams in the tests, with p-values less than or equal to 0.05. This study went on to show that there was a decrease in the phenotypes of PPD when Omega-3 fatty acids were incorporated, suggesting that this could be applied to human mothers as well (Gonzales et. al. 2015).

The current methods of treating PPD revolve around anti-depressants and therapy. Nutrition has not been explored as deeply for potential mechanisms for treating PPD. However, with the research seen with rodents, and with longitudinal studies done with women with and without PPD, there has been an increase in research for this field. The primary way to incorporate more Omega-3 into diets is around taking pills or more fatty acid enriched foods on a regular basis. But a more convenient way that has already been explored is a transdermal method of introducing the nutrients.

The bodily systems of mothers in their postpartum period are often abnormal due to the loss of energy and hormone levels. One of the main systems that are affected is the digestive systems. Incorporating pills or other food items could possibly cause more discomfort. Transdermal patches have been used for hormonal therapies that have also aided PPD recovery, showing that this mechanism could be extrapolated for this use. However, one main drawback of transdermal patches is that they have set quantities that are purchased online, and the dosages currently available on the market are too low for the amount needed to sustain recovery for mothers with PPD. Thus, another transdermal mechanism that hasn't been explored can be utilized for the correct dosages: microneedle patches. Microneedle patches haven't been designed for this purpose yet, and haven't been introduced to the nutrition-based payloads for use either, so creating a patch with Omega-3 fatty acids could serve as a tool to help mothers with PPD.

Utilizing a controlled delivery system of these fatty acids via a patch would prove revolutionary in helping mothers with PPD control

their symptoms. Though there has already been research on taking Omega-3 supplements and incorporating it into mothers' diets, no research has been done on the other various delivery methods, including transdermal microneedle patches.

Transdermal microneedle patches have been proven to be one of the most effective ways to deliver medication and hormonal therapies. Using this knowledge, a microneedle patch could be utilized to provide the large doses needed for mothers with postpartum depression in a convenient form that is easier on the physical level. The final product, O-Patch, should produce 1-2 grams of Omega-3 fatty acids in a simulated skin model over a period of 24 hours, successfully diffuse the payload solution to a depth of at least 350-450 micrometers, and have a steady release rate of at least 0.063 grams per hour.

METHODS AND MATERIALS

Materials

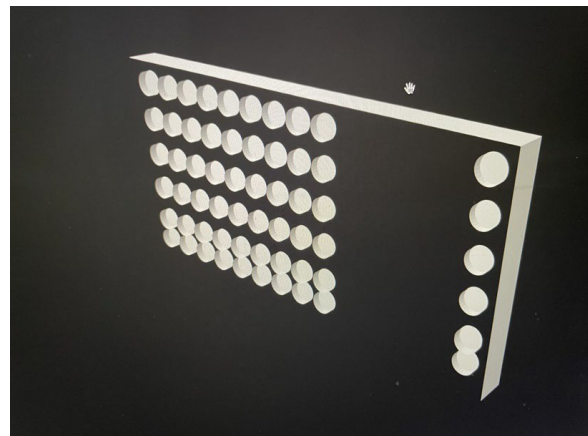
The materials that were used in this study included 1/16th of a cup of Polyethylene Glycol (PEG) and 1 tablespoon of KNOX Gelatin. These two substances were used for the base patch. 2 tablespoons of saline solution (0.5% sodium chloride), 50 mL of Hyaluronic Acid, 2 drops of red food dye and 3 Omega-3 Pills (Nordic Naturals) were used to create the payload solution. A rectangular mold, which was 30 millimeters by 60 millimeters was used for the base of the patch for size measurements. A 10 percent gelatin solution and parafilm from Parafilm Product were used for the simulated skin model that was used for testing. ImageJ was used to analyze the depth in micrometers and the release rate over a period of time. 3 glass dishes were used for testing and as containers for the simulated skin model. Latex

gloves and eyewear goggles were utilized as proper protective equipment while working with the patches and testing.

Patch Fabrication

Firstly, mold creation was initiated. The 3-D printer Tesla STEM High School was utilized, and a CAD design on FreeCAD served as the model. This included the needles jutting out at a 1000-micrometer depth with a flat base, and the needles were hollow to a diameter of 3 millimeters for medium-sized microneedles. The final mold with these dimensions are seen in Figure 1.

Figure 1:



CAD of Patch Mold

This configuration of 60 wells, with 3mm x 1mm dimensions, allows for a proper release rate with the appropriate amount, not too little or much, being released continuously.

Next, the payload solution was created. 2 tablespoons of saline solution (0.5% sodium chloride), 50 mL of Hyaluronic Acid, 2 drops of red food dye and 3 Omega-3 Pills of 500 mg each were mixed at 25 degrees Celsius and was set aside until the base solution was made.

This process is continued with microneedle fabrication. 1/16 of a cup of PEG and 1 tablespoon

of gelatin were homogenized at 55 degrees Celsius. This creates a cross-linkage between the two materials and allows for a matrix in which the Omega-3 fatty acids can sit and diffuse from. Then, the mixture was cooled to 20 degrees Celsius before the payload solution was combined in and stirred evenly for 2 minutes. Then, this solution was poured into the mold and was left to solidify at 5 degrees Celsius. This process took 15 hours, after which the patch was removed for testing. This fabrication process was repeated three times to create three patches.

Simulated Skin Model

A simulated skin model was developed following the procedure from Sadeqi, A., Kiaee, G., Zeng, W., Rezaei Nejad, H., & Sonkusale, S. A 10% gelatin-water solution was created in a small glass container. The solution was then poured into a flat petri dish with a depth of at least 0.5 cm for proper depth analysis. The gelatin-water solution should be dry halfway (about 65 minutes) before carefully adding parafilm to the top of the solution. The parafilm was draped to cover the entire surface of the plate. This setup was left to dry completely for about 2-3 hours at 30 degrees Fahrenheit.

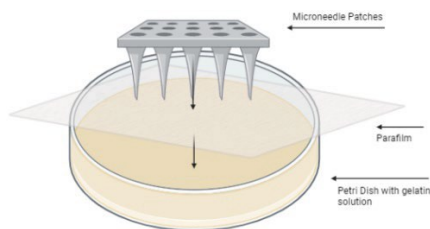


Figure 2: Diagram of Skin Model

The stratum corneum (uppermost waterproof layer of skin) is simulated by the parafilm and the epidermis section would be the gelatin solution. The insertion technique here shows the similarity and reasoning for using the simulated model, as it very closely replicates the human skin model.

Testing

The microneedle patch was then applied to the membrane by slowly placing it onto a clean, smooth portion of the simulated skin model, mimicking the application process on human skin. After 24 hours of the patch sitting on the simulated skin model, high-resolution pictures were captured using the cameras on the iPhone 12 and fed into ImageJ in TIFF format to obtain quantitative depth measurements in micrometers. These images were taken of the penetration of the payload solution into the model. This process is represented by Figure 3.

Specifically, in the browser version of ImageJ, the "Measurements" option was selected, displaying the length of the divot in micrometers. First, a reference length of 60 mm was set to calibrate the ImageJ software. Next, 20 representative points along the length of the patch were measured and taken as an average for penetration data. The results of this image analysis were recorded in Table 1.1 for further evaluation and comparison.

Table 1.1

Patches	Penetration (mm)
Patch 1	1.466
Patch 2	1.224
Patch 3	1.287

Set-up for Depth Analysis

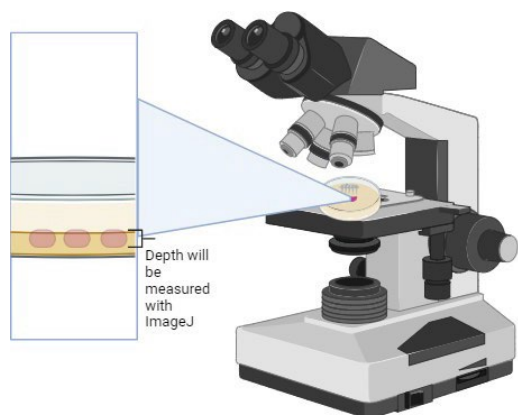


Figure 3 The pictures and the measurements will be taken in micrometers following this set-up. The colored fluid, as depicted here, will provide ample depth to analyze using the ImageJ software.

Using this penetration data, the release rate and total amount released were calculated using the following equation:

$$9 \pi \frac{\text{mmmm}}{\text{mmmm}} \cdot 60 \cdot (PPPP - PPPP) = 1000 \cdot TTTTTT$$

The $9\pi\pi$ mg/mm is the volume of each well, multiplied with the 60 wells from the mold, which is

then multiplied with the final penetration (Pf) minus the initial penetration (Pi). This was then divided by 1000 to get the total value released in grams, and the release rate was then calculated by dividing by 24 to get the grams released per hour.

The calculated rate of release was recorded in Table 1.2 for documentation and analysis.

Table 1.2

Patches	Release Rate (g/hr)
Patch 1	0.104
Patch 2	0.09
Patch 3	0.091

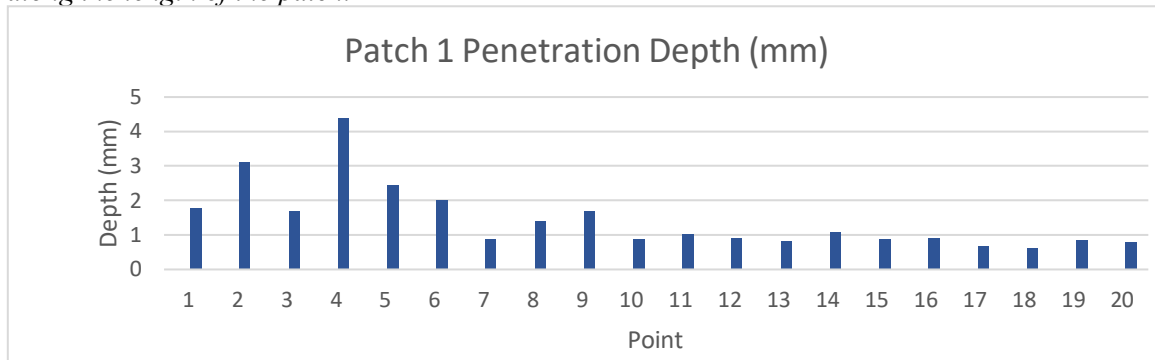
The total amount released per patch is recorded in this data table for analysis.

Table 1.3

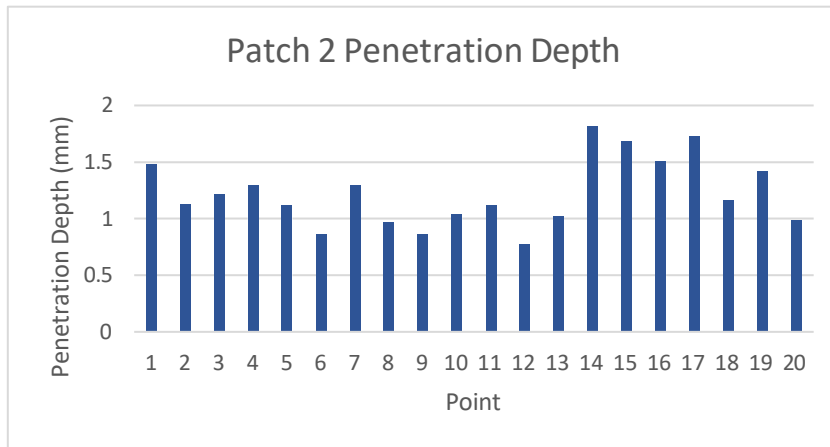
Patches	Total Amount (grams)
Patch 1	2.49
Patch 2	2.08
Patch 3	2.18

RESULTS:

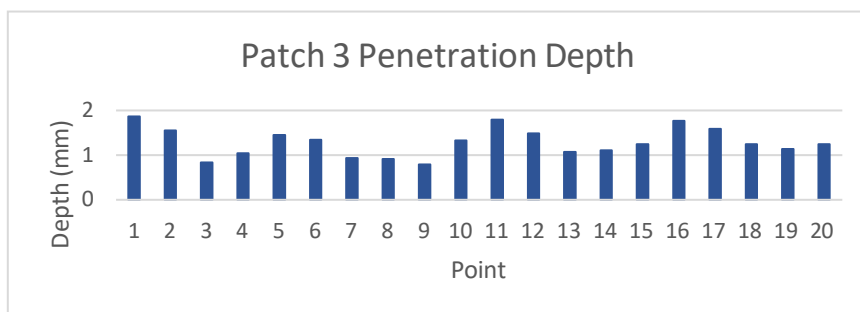
Three patches were created, and the graphs below indicate the penetration depths measured across 20 points along the length of the patch.



Graph 1: Shows the penetrations depth along 20 random points of the first patch in the simulated skin model, with 2 major outliers.



Graph 2: This shows the penetration depth for 20 random points across the second patch, with 2 outliers.



Graph 3 This shows the penetration depth data for 20 representative random points along the third patch, with 2 outliers.

The standard deviation for Patch 1 was significantly higher than the deviations for Patch 2

and Patch 3, due to the wells being artificially created first. The standard deviations were 0.98, 0.29, and 0.32 respectively for Patches 1, 2 and 3.

There were a few exiguous outliers, as seen by Point 4 in Patch 1, Point 1 and Point 11 in Patch 3, and Points 14 and 17 in Patch 2. These outliers are most likely due to air bubbles in the simulated skin model. When the microneedle patches penetrate the skin model, several of these air pockets are filled up with the payload solution, which illustrates a larger penetration depth than would be normally expected. To mitigate this, in future studies the skin model should be vigorously tapped in its liquid stage to release all air bubbles.

For patch 1, the mean penetration depth was 1.466 mm, the release rate of the solution was 0.104 g/hr, and the total amount that was released was 2.49 grams. For patch 2, the mean penetration depth was 1.224 mm, and the release rate was 0.09 g/hr with a total amount of 2.08 grams being released. Finally, for Patch 3, the average penetration depth was 1.287 mm, and there was a release rate of 0.091 g/hr with a total amount of 2.48 grams.

In total, these three patches account for three trials, resulting in a mean penetration depth of 1.319 mm, an average release rate of 0.095 g/hr and a total amount of 2.25 grams of solution released on average. These values significantly exceed the engineering goals set as well as current industry standards.

Using a Student's T-test, t-scores were calculated for the three variables.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

T = T-test score, μ = sample mean, s = standard deviation, n = sample size, x = population mean

Using this formula, where the population mean is the engineering goals and industry standards, the sample size is three, and sample means were the calculated averages for the variables, t-scores of 15.8 (for penetration depth), 7.23 (total amount released), and 6.9 (release rate) were calculated. These scores show there was a large variance and deviation from the population mean compared to the sample mean, showing significance. To further attest to this, p-values were created in Excel from these t-test scores.

This equation was used:

`TT.TTTTTTTT(aaaaaaaaaaaaaa, aaaaaaaaaaaaaa, TTaaTTTTTT, TTaaTTTT)`

With array1 and array2 holding the industry standards and sample means respectively, and a tail and type of 1 and 1 respectively, the following p-values were calculated: 0.0037 for penetration depth, 0.018 for total amount released and 0.02 for release rate. A 1 for Tails indicates it's a 1-tailed distribution, whereas the 1 for the Type entails a paired t-test. These values are statistically significant since they are less than 0.05. This shows that the base patch created significantly exceeds current industry standards and the engineering goals set, allowing for efficient diffusion of the Omega-3 payload solution into the skin.

CONCLUSIONS

All the engineering goals were exceeded significantly, showing success of the patch in being used to diffuse Omega-3 fatty acids. The average values for penetration depth, release rate and total amount were significantly higher than the goals set,

which is shown by the statistical significance in difference between the two determined by the t-scores and p-values. The unit cost per patch is \$15 with integrated microneedles made from laser ablation and microprinting, which is significantly less than the prices for microneedles patches on the market currently. With a significant deviation from industry standards in terms of efficiency, affordability, and payload amount, this patch can be continually improved and used to treat postpartum depression in tandem with therapies and medication.

Future Implications and Steps

In rural communities and urban areas, the diagnosis of postpartum depression is growing significantly, which can be combated in part with the integration of the patch during the postpartum period and 3 months before delivery. Mathematically modeling this, conclusions from several research articles suggest that an improved O-Patch can decrease the severity and symptoms of about 83% of current postpartum cases (Hoge et al. 2019). With the participants in this study, it was found that at a higher intake of Omega-3 fatty acids, approximately 5% of the daily caloric minimum, 83% of the population tested received a number less than 13 on the Edinburgh Postnatal Depression Scale, indicating an extremely low chance of postpartum depression. This means that from the 2100 caloric intake minimum for women, 105 calories—or about 10 grams of Omega-3 fatty acids—can be ingested to reduce the risk of postpartum depression by 83%. Annually, there are 4 million live births and several still-births and miscarriages that contribute to the birth count. Out of this number, about 900,000 mothers will be diagnosed with postpartum depression, but with this patch improved to a higher

release amount, the severity and symptoms of 749,700 cases will decrease significantly.

With the current O-Patch results, if the ratio of 10 grams to 83.3% reduced risk is constant, then with 2.5 grams being released currently, 27% of cases can be reduced in severity. With 900,000 annual cases, this percentage represents 243,097 mothers who would have more manageable and lessened symptoms of postpartum depression.

To reach the possibility of nearly 83% of cases being decreased in severity, continual testing of PEG/Gelatin ratios and patch sizes must be conducted. Large microneedle patches can have surface area dimensions as large as 20 cm by 5 cm, which would greatly increase the amount of payload solution released. Increasing the size of the patch during testing can show a larger amount being released into the simulated skin model and will release at least 5 to 7 grams more than current amounts. Additionally, possibly increasing the PEG to gelatin ratio will allow for Omega-3 fatty acids to diffuse into the skin at a faster rate. However, this method will require biocompatibility testing to make sure the increased rate of defusal doesn't negatively affect the nutritional homeostatic states of the body and cause negative side effects like increased cholesterol levels.

Furthermore, concentrations of the gelatin and PEG can be altered for different payload substances to deliver various solution in different medical and industrial applications, such as hormonal therapies, medicinal delivery systems, and nutritional delivery systems. Because this patch has been created newly, the payload solutions can vary, allowing for a wide variety of use cases beyond just Omega-3 defusal for postpartum depression. Some

specific examples include nicotine patches to aid addiction recovery, or estrogen hormonal therapies. This patch can be used in the future after

biocompatibility testing and reliability testing by way of twenty or more patches being tested in vitro and through clinical testing.

BIBLIOGRAPHY

- da Rocha, C. M., & Kac, G. (2012). High dietary ratio of omega-6 to omega-3 polyunsaturated acids during pregnancy and prevalence of post-partum depression. *Maternal & child nutrition*, 8(1), 36–48. <https://doi.org/10.1111/j.1740-8709.2010.00256.x>
- De Vriese, S. R., Christophe, A. B., & Maes, M. (2003). Lowered serum n-3 polyunsaturated fatty acid (PUFA) levels predict the occurrence of postpartum depression: further evidence that lowered n-PUFAs are related to major depression. *Life sciences*, 73(25), 3181–3187. <https://doi.org/10.1016/j.lfs.2003.02.001>
- Faraji Rad, Z., Prewett, P. D., & Davies, G. J. (2021, September 13). *An overview of microneedle applications, materials, and fabrication methods*. Beilstein journal of nanotechnology. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8450954/>
- Gonzales, E., Barrett, D. W., Shumake, J., Gonzalez-Lima, F., & Lane, M. A. (2015). Omega-3 fatty acids improve behavioral coping to stress in Multiparous rats. *Behavioural Brain Research*, 279, 129–138. <https://doi.org/10.1016/j.bbr.2014.11.010>
- Hoge, Tabar, Donneau, Dardenne, Degée, Timmermans, Nisolle, Guillaume, & Castronovo. (2019). Imbalance between Omega-6 and Omega-3 Polyunsaturated Fatty Acids in Early Pregnancy Is Predictive of Postpartum Depression in a Belgian Cohort. *Nutrients*, 11(4), 876. <https://doi.org/10.3390/nu11040876>
- Oh, N. G., Hwang, S. Y., & Na, Y. H. (2022, July 14). *Fabrication of a PVA-based hydrogel Microneedle Patch*. ACS Publications. <https://pubs.acs.org/doi/10.1021/acsomega.2c01993>
- Ranamukhaarachchi, S. A., Schneider, T., Lehnert, S., Sprenger, L., Campbell, J. R., Mansoor, I., Lai, J. C., Rai, K., Dutz, J., Häfeli, U. O., & Stoeber, B. (2015). Development and validation of an artificial mechanical skin model for the study of interactions between skin and microneedles. *Macromolecular Materials and Engineering*, 301(3), 306–314. <https://doi.org/10.1002/mame.201500320>
- Sadeqi, A., Kiaee, G., Zeng, W., Rezaei Nejad, H., & Sonkusale, S. (2022). Hard polymeric porous microneedles on stretchable substrate for transdermal drug delivery. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-05912-6>
- Vernier. (2014). (*Order code GPS-BTA*) - vernier. <https://www.vernier.com/files/manuals/gps-bta/gps-bta.pdf>

ENVIRONMENTAL ENGINEERING

Effects of Pharmaceuticals and Ocean Warming on Zooxanthellae Expulsion and Asexual Reproduction in *Aiptasia*: A Model for Tropical Coral

Sanjana Katti

ABSTRACT

In the last ten years, coral reef surface area has decreased by nearly 30% due to bleaching and organismal death, causing increased storm damage to coastal cities, decreased support for marine ecosystems, and economic loss from tourism. Among possible stressors for this loss are ocean warming (led by climate change) and pharmaceutical products, which are emerging concerns for their potential endocrine-disrupting effects. In this research, an experiment was conducted using *Aiptasia*, a model organism for coral, and measured reproductive harm and bleaching caused by heated ocean temperatures and pharmaceutical pollution concentrations. Eight aquaria were set up with varying water temperatures and acetaminophen/ibuprofen concentrations. Every 48 hours, the *Aiptasia* were counted and the zooxanthellae algae cells, expelled during bleaching, were filtered to count using a hemocytometer. High zooxanthellae cell expulsion indicated high thermochemical stress in the *Aiptasia*, and low *Aiptasia* populations indicated reproductive inhibition by either or both temperature and pharmaceutical concentration. The results show that a medium concentration of pharmaceuticals (2.1 µg/L acetaminophen and 2.4 µg/L ibuprofen) at 27°C caused the greatest change in *Aiptasia* population relative to the control, with a p-value of 0.00064, and that population was significantly affected by temperature for all pharmaceutical concentrations ($p < 0.002$). Zooxanthellae expulsion was not significantly affected by either pharmaceutical concentration or temperature ($p > 0.06$), necessitating additional investigation. This study highlights a need for stricter pharmaceutical regulations in marine ecosystems and further research into coral physiology and resilience to prepare for a warming ocean and its global consequences.

INTRODUCTION

The objective of this experiment is to suggest the effects of common over-the-counter medications and increased ocean water temperature (due to climate change) on coral reproduction and bleaching.

Pharmaceutical compounds in marine ecosystems

Pharmaceutical compounds are contaminants of growing environmental concern, situated within the broader category of Pharmaceuticals and Personal Care Products (PPCPs). The United Nations Environment Program

(UNEP) notes that pharmaceuticals are “biologically active substances intended to cause pharmacological effects in living organisms,” and as such, they also have an impact on wildlife and the ecosystem health when improperly treated (UNEP, 2020). In fact, several pharmaceuticals such as ibuprofen, diclofenac, naproxen, and acetaminophen have been detected in concentrations upwards of 3.35µg/L in wastewater effluent. Effluent is sewage that has been treated by a wastewater facility and flows directly into surface waters such as lakes, rivers, and oceans (Zanuri, 2016). Although most pharmaceuticals are

detected in the nanogram-per-liter (ng/L) to low microgram-per-liter ($\mu\text{g/L}$) ranges, the risks to aquatic life cannot be ignored due to the bioactive capacity of pharmaceuticals. Notably, sperm exposure to just $1.0\mu\text{g/L}$ of ibuprofen affects the fertilization success and gamete production of two marine invertebrates, *P. miliaris* (a species of sea urchin) and *A. marina*, a polychaete worm (Zanuri, 2016). These pharmaceutical remnants are of paramount concern because seawater hosts one of the most fragile yet diverse ecosystems on the planet: coral reefs.

Coral physiology

Coral reefs teem with life, providing habitat and food to over 30% of all marine organisms while occupying just 0.2% of the sea floor (Woodall, 2020). They are even called “the rainforests of the sea” for their immense species diversity, especially in tropical waters such as those in the Malay Archipelago, Australia, the Pacific Islands, and the Caribbean. Coral reefs are formed over millions of years and are made up of thousands or millions of individual coral polyps. These polyps are not just the structural basis of the reef’s ecosystem, but also bioindicators of climate change due to their sensitivity to changing ocean temperature and pH. This sensitivity arises from the coral’s symbiotic relationship with photosynthetic algae called zooxanthellae, which produce over 90% of the food that the coral requires to produce bright colors and calcify (form skeletons). Once the coral expels its zooxanthellae due to thermochemical stress, the organism loses not just its color, but its primary source of nutrients in warm and generally nutrient-poor water. The result of this process, called bleaching, culminates in coral polyps starving. As

warm temperatures produce a similar effect across the reef and more polyps die, the reef struggles to reproduce and grow, so fewer fish are able to use the reef as their home, causing harmful algae to build up without an organism to feed on it. The buildup of algae causes the coral polyps to suffocate, and the reef dies. The implications of this are immense, as coral reefs are natural barriers to waves and storm surges, absorbing the strength of the water to protect coastal communities. Furthermore, bleached corals deprive fish and crustacean species of a habitat to spawn and develop, compounding the ecological stress. It is for these reasons that in 2020, the United Nations Environmental Program named coral as the organism having the most rapid increase in extinction risk out of any species cataloged. As such, it remains a concern that coral bleaching at a mass scale has increased substantially in recent years, both in the area affected and frequency of occurrence. Several studies by NASA and NOAA have shown that rising ocean temperatures are a primary cause for bleaching events, but other possible factors include ocean acidification and anthropogenic pollution, including pharmaceuticals.

Drugs studied in past and present research

Non-steroidal anti-inflammatory drugs (NSAIDs) are the most widely consumed category of pharmaceutical drugs (Fent et al., 2006), used primarily for pain and inflammation relief. In this study, ibuprofen is used for toxicological experiments as it is water soluble, bioaccumulative, and persistent, indicating a potential presence in sea environments and the organisms that reside in them (Bendz et al., 2005, Maurer et al., 2007). This study also uses acetaminophen, a non-NSAID analgesic bought over the counter. The combination of

acetaminophen and ibuprofen has a more powerful effect than either pharmaceutical alone, indicating that their combined exposure will likely have an additive effect on studied organisms (Harvard Health, 2011). Currently available studies on acetaminophen toxicity on any marine invertebrate are sparse, limited to a study on the edible clams *V. decussata* and *V. philippinarum* Antunes et al., 2013), neither of which are closely related to coral. Additionally, although ibuprofen toxicity in invertebrates has been studied, including on the species *Daphnia magna* (Heckmann et al., 2007), no studies have been conducted on coral to date. Ultimately, the lack of research on coral translates to an incomplete risk assessment, demonstrating a need for further discovery on this topic.

Experimental concessions

One notable limitation of coral use in scientific studies, however, is that field collection of coral is destructive, and seasonal reproduction cycles and a series of protection laws limit collection, even for scientific studies. However, the organism *Aiptasia*, a species of sea anemone, can be used as a photosynthetic and metabolic model for coral. *Aiptasia* and coral belong to the same phylum (Cnidaria) and have similar physiology, so *Aiptasia* are regularly used as a model for coral physiology in scientific research. This experiment hypothesizes that 1) pharmaceutical pollution will have a greater effect on reproduction than zooxanthellae expulsion, and 2) there will be a measurable difference between the zooxanthellae expelled in high pharmaceutical concentrations against the control.

MATERIALS AND METHODOLOGY

Figure 1: Simplified Procedure

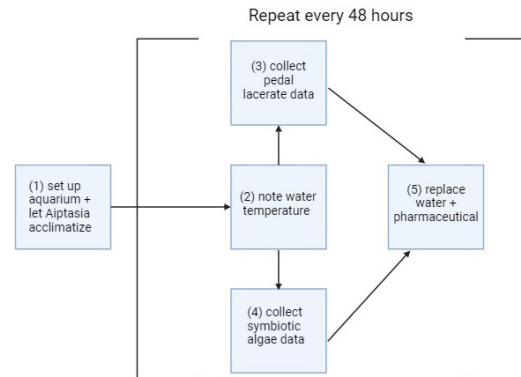


Figure 1: Graphical summary of experimental procedure: setup, collection of zooxanthellae, water, and pedal lacerates, and replacement of water and pharmaceutical concentration.

Treatment Materials

For this study, a sample of 95 reproductive *Aiptasia* was acquired from Seattle Corals. A bag of limestone rock was obtained from Saltwater City, and each rock was broken into 3-7 cm pieces with a hammer. Each rock contained seeded bacteria and was cured for 6 days in $27 \pm 1^\circ\text{C}$ saltwater from Saltwater City to activate the bacteria. Advil Liqui-Gels minis, with 200mg of solubilized ibuprofen, was bought from Amazon, and Children's Tylenol, with 160mg/5mL of acetaminophen, was bought from Amazon.

Aquarium Setup

Eight clear plastic boxes approximately 3.75L in volume were acquired from IKEA; these were the aquaria. 8 water heaters and 8 air pumps were bought, one of each for every aquarium. Approximately 2L of seawater, acquired from Saltwater City, was poured into each aquarium, such that both the water heater and air pump in each tank

were completely submerged. 5 rocks were placed into each aquarium, approximately equidistant from each other. The pump was turned on, and the water heater was set to $27 \pm 1^\circ\text{C}$. Once the water had been warmed, 2 *Aiptasia* were placed on each rock using rubber gloves. A plastic cover was partially placed over the aquarium to discourage evaporation and maintain water temperature, and a SEAOURA LED Aquarium Light, giving 12 continuous hours of blue light to simulate daylight and 12 hours of mixed-color light to simulate nighttime, was placed over the 8 aquaria.

Figure 2: Experimental Setup



Figure 2: Each aquarium contains 10 *Aiptasia* anchored to rocks, 1 heater, and 1 air pump. Rocks omitted for simplicity.

Figure 3: Photograph of Experimental Setup

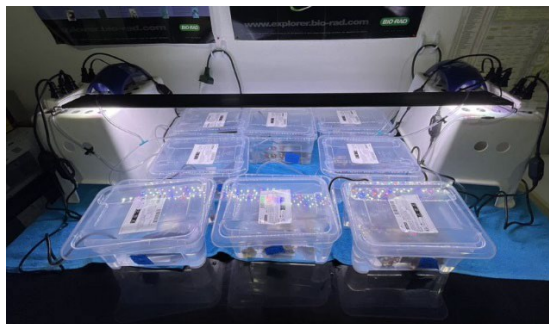


Figure 3: Top view of 8 aquaria, placed equally under 1 SEAOURA LED light for the course of the experiment.

Pharmaceutical Solutions

200mg solubilized ibuprofen and 160mg acetaminophen were diluted with seawater to form the following concentrations (*Table 1*). 2 aquaria (the control tanks) did not have pharmaceuticals.

Table 1: Concentrations of Ibuprofen and Acetaminophen

Solubilized Ibuprofen	Acetaminophen
0.8 $\mu\text{g/L}$	0.7 $\mu\text{g/L}$
2.4 $\mu\text{g/L}$	2.1 $\mu\text{g/L}$
6.4 $\mu\text{g/L}$	4.9 $\mu\text{g/L}$

Table 1: Sorted concentrations of pharmaceuticals for use in the experiment.

These concentrations were devised as the range of ibuprofen and acetaminophen is extremely large, so concentration ranges from 19 scientific articles studying 5 locations were averaged to find a low end of detection, average end, and high end of detection.

Data Collection

Every 48 hours for 15 days, 2 types of data were collected: number of pedal lacerates (asexually produced offspring) and quantity of zooxanthellae expelled. To quantify the pedal lacerates, a magnifying glass was used to inspect each *Aiptasia* specimen for possible pedal lacerates. A photo was taken for further analysis. Data was stored in *Table 3*.

Table 3: Number of Pedal Lacerates Over Time

	Day 1	Day 3	Day 5	Day 7	Day 9	Day 11	Day 13	Day 15
27°C L	11	11	11	12	13	11	15	14
30°C L	11	14.3	9.9	2.2	2.2	2.2	2.2	2.2

To quantify the zooxanthellae expelled, rocks with Aiptasia attached were removed from test container and placed in a temporary container with seawater. All seawater in the test container was poured through 100% cotton cloth, then

zooxanthellae cells were scraped off cotton cloth with a razor blade, into a vial with 3mL seawater. The cylinder was shaken and 0.05mL was placed onto a hemocytometer grid. Zooxanthellae cells were counted in 5 of 25 measurement areas (upper left & right, lower right & left, center). A photo was taken of the hemocytometer grid for future crosschecking analysis. The number of cells in each grid were averaged, and the calculated value was converted to cells/mL/specimen using the following formula (www.jstor.org/stable/3628384). Data was stored in *Table 4*.

$$\frac{\text{MMMMMMMM hMMeeeeeeeeeeeeMMeeMMee eeeccMMee} * 50000 \text{ hMMeeeeeeeeeeeeMMeeMMee MMaaaccacaeMMee} * 3\text{eemm aadddcceeddeeMM}}{2\text{mm MMAaccMMeeddccee vveeddcceeMM} * 10 \text{ aassMMeeddeeMMMMaa}}$$

Table 4: Zooxanthellae Loss Over Time (cells/mL/specimen)

	Day 5	Day 7	Day 9	Day 11	Day 13	Day 15
27°C L	118	127	34	73	8	62
30°C L	192	29	0	9	1	16

Table 4: Units are cells/mL/specimen and days 1 and 3 were not stored because samples were contaminated during experimentation.

Data Analysis

The number of Aiptasia per aquarium at the start of the experiment was variable due to the anchoring tendencies of Aiptasia, and as such, the starting count was scaled to 11 during data analysis for each control and treatment aquarium, and the proceeding values were also scaled as needed. These values were graphed in Excel, and the outliers

circled. A one-way ANOVA was used to compare the variances between experimental groups, as well as between control and treatment groups. The below circled outlier was explained as due to human error, but the significance regardless indicates a statistically significant difference between the treatment and control tanks.

Figure 3: Aiptasia Population Over Time, 30°C & High Concentration

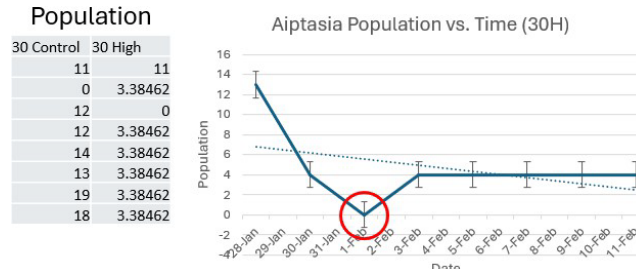


Figure 3: Comparison of Aiptasia population between 30°C tanks at a control and high concentration of pharmaceuticals. The calculated p-value was 0.001662.

Figure 4: Aiptasia Populations Over Time - Comparison

30 Control	30 Medium	30 Control	30 Low	27 Medium	30 Medium	27 High	30 High
11	11	11	11	11	11	11	11
0	11.733333	0	14.3	13.2	11.73333	11.91667	3.384615
12	7.333333	12.00001	9.9	14.66667	7.333333	11.91667	0
12	1.466667	12.00001	2.2	15.4	1.466667	11.91667	3.384615
14	2.2	14.00003	2.2	16.13333	2.2	9.166667	3.384615
13	1.466667	13.00002	2.2	16.86667	1.466667	9.166667	3.384615
19	2.2	18.99997	2.2	15.4	2.2	10.08333	3.384615
18	4.4	17.99996	2.2	16.13333	4.4	11.91667	3.384615

Figure 4: Two aquaria were compared: 30°C at a control and medium concentration of pharmaceuticals, 30°C control and low concentration, 27°C and 30°C medium concentration, and 27°C and 30°C high concentration. The calculated p-values for all four comparisons were <0.05.

RESULTS

The experiments were conducted every 48 hours for 15 days, and group comparisons were conducted using one-way analysis of variance (ANOVA) via the Excel extension. Graphs were generated using Excel, with each graph displaying a data point $\pm$ standard deviation (SD) and trendline. It was concluded that a statistically significant relationship exists between both temperature and pharmaceutical concentration on Aiptasia reproduction ($p < 0.02$ for all but 27°C control/27°C low and 27°C control/27°C high comparisons). Outliers were critically analyzed; outliers took the form of an abnormal drop in population or an uncharacteristic rise in population. At 30°C and a high pharmaceutical concentration, displayed below,

an uncharacteristic drop in population was observed on day 3 of experimentation – from 4 to 0 Aiptasia.

Figure 5: Aiptasia Population vs. Time at 30°C & High Concentration

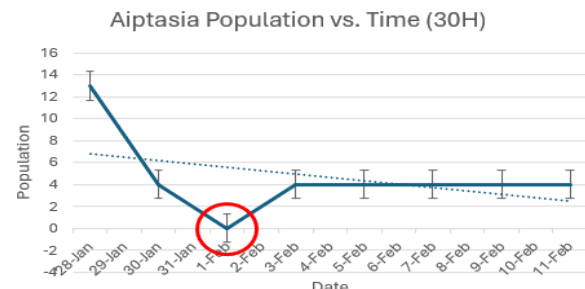


Figure 5: Aiptasia population at 30°C at a high concentration of pharmaceuticals displayed a downwards trend. One outlier on the third day of experimentation is circled in red.

As Aiptasia populations were counted by hand, it is likely that the 4 missing Aiptasia were present on day 3 but were miscounted. Newly reproduced Aiptasia are typically 2-5mm in length

and nearly transparent, making them very difficult to observe. Seeing as the population stabilized after this drop, the likelihood of this being the reason for the outlier is high. A similar observation was seen at 27°C and a low pharmaceutical concentration.

However, another outlier lay in 6 of the 8 zooxanthellae graphs. On day 2 of experimentation, a large spike in expelled zooxanthellae was observed.

Figure 6: Expelled Zooxanthellae vs. Time at 30°C & Medium Concentration

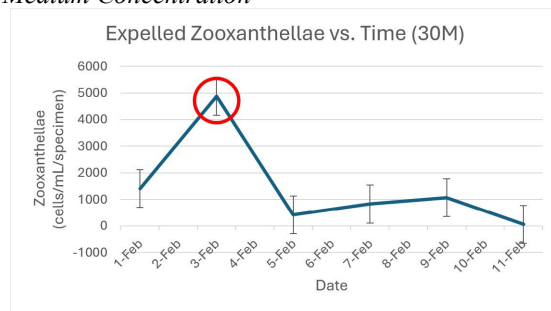


Figure 6: Graph of Aiptasia population at 30°C at a medium concentration of pharmaceuticals. One outlier on day two of experimentation is circled in red.

As is visualized in the above graph, expelled zooxanthellae cells were consistent besides the outlier on day 2 of experimentation. One possible explanation for this outcome is that the Aiptasia were under great thermodynamic stress at 30°C, having been moved from a holding tank at 27°C just four days earlier. It is very likely that in response to this stress, the Aiptasia did in fact expel an increased concentration of zooxanthellae cells, but as the species acclimatized to the environment, this number dropped to a more consistent value. It is important to note that the average values without the outliers for the 27°C control is largely different from the 30°C control, but not enough to have statistical significance.

DISCUSSION

A statistically significant relationship was found between both pharmaceuticals and temperature on Aiptasia reproduction. No significant relationship was found between either variable on zooxanthellae expulsion – the latter is an incorrect finding. It is a well-established phenomenon, as per the United Nations Environmental Programme in 2021, that increased water temperature is a significant cause of coral bleaching. As such, several theories were developed as to the possibility of experimental error. First, cotton is an absorbent material. During the filtration of aquarium water, a possibility remains that the cotton absorbed some zooxanthellae cells, making it difficult to scrape off and into a vial for counting. Second, that the data was miscounted. Under the hemocytometer, several types of cells were visible, and zooxanthellae clusters were difficult to distinguish from additional organisms. The likelihood of a miscount is extremely high, but a digital method of counting could decrease this risk. Third, that during the filtration process, some algae were tethered to rocks and the heater present in the aquarium. This implies that although the algae cells were expelled, they were not entirely removed from the container.

The Endocrine Disruption Hypothesis was first proposed from an ecological standpoint by Dr. Theo Colborn as part of his 1996 book *Our Stolen Future*. Colborn raised alarm to the potential risks caused by endocrine disrupting compounds, including pharmaceuticals, to wildlife. The hypothesis states that these compounds mimic or block hormones and can lead to reproductive harm,

most notably infertility. The findings presented in this experiment support Colborn's hypothesis: even minute exposure to pharmaceutical pollution (2.1µg/L and 2.4µg/L) had a significant effect on Aiptasia reproduction, with a p-value of 0.00244 when compared to the control. This phenomenon can be explained because these pharmaceuticals are more potent to Aiptasia and coral polyps than human beings, as these organisms are orders of magnitude smaller and less complex than humans. As such, a minute dose to humans still has the potential to cause significant reproductive harm to these organisms, as was demonstrated throughout the experiment.

Future research on this topic includes isolating each Aiptasia, rather than placing several in one tank. In this way, several iterations can be made within each temperature and pharmaceutical concentration, and a stronger correlation can be made with zooxanthellae expulsion as each Aiptasia expels zooxanthellae into a separate container. Additionally, the cellular mechanisms behind mitosis and photosynthesis should be examined to determine the molecular effects of pharmaceuticals on Aiptasia physiology. This research is applicable to equatorial coral – found in the Malay Archipelago, the Caribbean, Australia, and the Red Sea. As the concentrations utilized in the experiment were environmentally relevant, the effects demonstrated in the experiment can be applied to these precise ecosystems. As such, several further steps and conclusions can be drawn: one, that

stronger regulation must be placed on pharmaceutical dumping and concentration limits in ocean water. Currently, current bans and regulations by the Environmental Protection Agency do not cover any type of pharmaceutical pollution. Furthermore, this study demonstrates the need for further protection of equatorial reefs through designation as protected areas. Ultimately, developing equatorial nations require support to build wastewater treatment infrastructure to prevent the passing of these pharmaceuticals into ocean water.

Coral reefs are among the most diverse yet fragile ecosystems on the planet and provide trillions of dollars to the global economy. As the ocean warms and pharmaceutical consumption increases, it becomes increasingly important to protect coral reefs and their surrounding ecosystems for the conservation of nature and the humans who depend on them.

ACKNOWLEDGMENTS

Foremost, thank you to Dr. Mark Ladd of the NOAA Southeast Fisheries Science Center for his patient guidance in the planning stages of the experiment. Thank you to SoundBio Laboratory for their gracious lab space given throughout the experiment, and to Mrs. Kate Allender of Tesla STEM High School for her continual support and guidance throughout the development and execution of this experiment. And finally, thank you to my parents for your cars and credit card. This project would not have been possible without you.

REFERENCES

- Antunes, S. C., Freitas, R., Figueira, E., Gonçalves, F., & Nunes, B. (2013). Biochemical effects of acetaminophen in aquatic species: Edible clams *Venerupis decussata* and *Venerupis philippinarum*. *Environmental Science and Pollution Research*, 20(9), 6658-6666. <https://doi.org/10.1007/s11356-013-1784-9>
- Bendz, D., Paxéus, N. A., Ginn, T. R., & Loge, F. J. (2005). Occurrence and fate of pharmaceutically active compounds in the environment, a case study: Høje river in Sweden. *Journal of Hazardous Materials*, 122(3), 195-204. <https://doi.org/10.1016/j.jhazmat.2005.03.012>
- Drugs in the water*. (2011, June 1). Harvard Health Publishing. https://www.health.harvard.edu/newsletter_article/drugs-in-the-water#:~:text=At%20this%20point%2C%20there%27s%20really%20no%20evidence%20of,unused%20medications%20at%20central%20locations%2C%20serve%20two%20purposes.
- Ericson, H., Thorsén, G., & Kumblad, L. (2010). Physiological effects of diclofenac, ibuprofen and propranolol on Baltic sea blue mussels. *Aquatic Toxicology*, 99(2), 223-231. <https://doi.org/10.1016/j.aquatox.2010.04.017>
- Fent, K., Weston, A., & Caminada, D. (2006). Ecotoxicology of human pharmaceuticals. *Aquatic Toxicology*, 76(2), 122-159. <https://doi.org/10.1016/j.aquatox.2005.09.009>
- Guse, A. (n.d.). *Aiptasia*. Guse Lab. https://guselab.de/aiptasia/?doing_wp_cron=1701409659.2419641017913818359375
- Heckmann, L., Callaghan, A., Hooper, H., Connon, R., Hutchinson, T., Maund, S., & Sibly, R. (2007). Chronic toxicity of ibuprofen to *Daphnia magna*: Effects on life history traits and population dynamics. *Toxicology Letters*, 172(3), 137-145. <https://doi.org/10.1016/j.toxlet.2007.06.001>
- Howe, P. L., Reichelt-Brushett, A. J., & Clark, M. W. (2014). Effects of Cd, Co, Cu, Ni and Zn on asexual reproduction and early development of the tropical sea anemone *Aiptasia pulchella*. *Ecotoxicology*, 23(9), 1593-1606. <https://doi.org/10.1007/s10646-014-1299-2>
- Maurer, M., Escher, B., Richle, P., Schaffner, C., & Alder, A. (2007). Elimination of β -blockers in sewage treatment plants. *Water Research*, 41(7), 1614-1622. <https://doi.org/10.1016/j.watres.2007.01.004>
- Microscopic anatomy*. (n.d.). Coral Disease and Health Consortium. <https://cdhc.noaa.gov/coral-biology/microscopic-anatomy/>
- Romanó de Orte, M., Clowez, S., & Caldeira, K. (2019). Response of bleached and symbiotic sea anemones to plastic microfiber exposure. *Environmental Pollution*, 249, 512-517. <https://doi.org/10.1016/j.envpol.2019.02.100>
- Siddiqui, S., & Bielmyer-Fraser, G. K. (2015). Responses of the sea anemone, *Aiptasia pallida*, to ocean acidification conditions and copper exposure. *Aquatic Toxicology*, 167, 228-239. <https://doi.org/10.1016/j.aquatox.2015.08.012>
- Status of coral reefs of the world: 2020*. (2021, October 5). International Coral Reef Initiative. <https://doi.org/10.59387/wotj9184>
- Summer, K., Reichelt-Brushett, A., & Howe, P. (2019). Toxicity of manganese to various life stages of selected marine cnidarian species. *Ecotoxicology and Environmental Safety*, 167, 83-94. <https://doi.org/10.1016/j.ecoenv.2018.09.116>
- Wissmann, S. M. (2003). The Effects of Elevated Ultraviolet B Radiation and Elevated Water Temperature on the Loss of Zooxanthellae from *Aiptasia pallida*. *Transactions of the Kansas Academy of Science*, 106(No. 1/2). <http://www.jstor.org/stable/3628384>
- Woodall, T. (2022, May 12). *Large-scale ocean sanctuaries could protect coral reefs from climate change*. Ohio State University. <https://news.osu.edu/large-scale-ocean-sanctuaries-could-protect-coral-reefs-from-climate-change/>
- Zanuri, N. B. M. (2016). *The effects of environmental contaminants and ocean acidification on reproductive success in marine invertebrates* [Doctoral dissertation, Newcastle University]. Newcastle University. <https://theses.ncl.ac.uk/jspui/bitstream/10443/3638/1/Mohd%20Zanuri%2c%20N.B.%202017.pdf>

Impact of Heat Events on Urban Heat Island Effects

Eva Jain

INTRODUCTION

The purpose of this project was to analyze temperature levels of cities of various urbanization levels in order to give a more holistic view of the effects of the extreme heat events and Urban Heat Island effects in these locations and the impacts of climate change.

Climate change has increasingly become a global concern, resulting in more severe storms, droughts, and most notably, hotter temperatures. As the effects of climate change become more prevalent, it has caused record setting temperatures worldwide. Around the world, the average temperatures in the months of June, July, and August 2023 measure 1.2 degrees Celsius warmer than the average summer from 1951-1980, and the 2023 summer was 0.23 Celsius warmer than any other summer on record (Fox et al., 2023). The United States Environmental Protection Agency (EPA) explains that these unprecedented temperature increases correlate with a rise in energy use, heat-related illness, and human mortality. According to the EPA, with this increase in temperature, there has also been an increase in heat wave frequency, duration, season length, and intensity. There used to be an average of 2 heat waves per year in the 1960s, which has increased to 6 per year in the 2010s and 2020s (EPA Heat Waves 2023). Temperature increases also correlate with a rise in air pollutants and decrease in air quality due to increased energy consumption for air conditioning systems and particulate matter released by wildfires, and these

negative impacts cause a plethora of human health problems such as respiratory conditions and an increase in lung cancer incidence (Grigorieva & Lukyanets, 2021).

Additionally, the world is becoming more urbanized. According to the United Nations, as of now, more than 1/2 of the world's population lives in urban areas, which is projected to increase to 2/3 by 2050. These urbanized cities replace natural land cover with surfaces that retain and absorb heat, such as concrete. Alternatively, rural cities have vegetation and bodies of water that help deflect radiation from sun and reduce temperature. Studies that have observed this increase in temperature in urban areas compared more rural areas have called it the urban heat island effect (UHI), which explains the tendency for urban areas to be warmer than surrounding areas with natural land cover (Tong et al., 2021).

But what has not been done is investigation of the correlation between extreme heat events and temperature in these rural, suburban, and urban cities in Denton County, Texas. This project seeks to investigate how extreme heat events in Denton County have influenced the urban heat island effects. Analysis was done to attempt to prove a positive correlation with heat wave event data from May - September of 2023 with the UHI in rural and urban areas in Denton, Texas. Pollution, temperature, and latitude and longitude (which can be used to be determined if a location is urban or rural) data comes from sensors (specifically PM sensor PA-II-SD)

deployed around Denton and is provided to this study by Dr. Lu Liang. Typically, studies have been done on a few weather stations in the city, but this data was from fine-grained remote sensors, which gave more accurate results. Publicly available data on temperature for classifying heat events is from the Texas Commission on Environmental Quality (TCEQ). The data shows a correlation of higher temperatures and pollution levels in urban areas during these extreme heat events and if these trends in data are significant, that means that the UHI effects are intensified by the heat events.

METHODOLOGY:

To complete this project, a Macbook Pro laptop (macOS Ventura Version 13.2, Apple M2 chip, 8GB RAM, 256GB hard disk) was used for computational power to develop and run models and Microsoft Excel was used to visualize, process, and analyze data.

The first part of the project involved pre-processing temperature data, specifically from Denton County, Texas over the summer of 2023 (May - September). This made it possible to observe the excessive heat events and their effect on UHI and UPI. To identify heat events during this time period, weather data was downloaded from the summer of 2023. To do this, data was collected from

the TCEQ (Texas Commission on Environmental Quality) website, specifically the Texas Air Monitoring System (TAMIS). In the TAMIS portal, a Raw Data Report (AQS) was started. After downloading data from the range of May 1, 2023

(inclusive) to October 1, 2023 (exclusive), data from counties, specifically Denton Airport South was extracted. The specific data points necessary were the Outdoor Temperature, Dew Point, and Relative Humidity parameters, including headings and measurements. A sample duration of an hour and file delimiter of a tab were chosen and then the report was created and downloaded.

When the report is downloaded, the report was opened in Excel. For every hour for each day, the heat index using the heat index formula was calculated. As defined by the National Weather Service (HI = Heat Index, T = Temperature, RH = Relative Humidity), the formula is

$$\begin{aligned} HHHH = & -42.379 + 2.04901523 * \\ & TT + 10.14333127 * RRHH - .22475541 * TT * \\ & RRHH - .00683783 * TT * TT - .05481717 * RRHH * \\ & RRHH + .00122874 * TT * TT * RRHH + .00085282 * \\ & TT * RRHH * RRHH - 0.00000199 * TT * TT * RRHH * RRHH \end{aligned}$$

If the RH is less than 13% and the temperature is between 80 and 112 degrees Fahrenheit, then the following adjustment is subtracted from HI:

$$\begin{array}{c} \begin{array}{c} 13 - \\ RRHH \end{array} \qquad \begin{array}{c} 17 - |TT - 95. | \\ \hline 17 \end{array} \\ \hline \begin{array}{c} 4 \end{array} \end{array} \quad \begin{array}{c} \text{AAAAAAATTAAAAATT} = \text{?} * \text{?} \end{array}$$

If the RH is greater than 85% and the temperature is between 80- and 87-degrees Fahrenheit, then the following adjustment is added to HI:

$$\begin{array}{c} \begin{array}{c} (RRHH - 85) \\ \hline 10 \end{array} \qquad \begin{array}{c} (87 - TT) \\ \hline 5 \end{array} \\ \hline \begin{array}{c} \text{AAAAAAATTAAAAATT} = \text{?} * \text{?} \end{array} \end{array}$$

If the resultant heat index value is below about 80 degrees Fahrenheit, a new formula is applied:

$$\begin{aligned} HHHH = & 0.5 * (TT + 61.0 + [(TT - 68.0) * 1.2] + (RRHH \\ & * 0.094)) \end{aligned}$$

Heat index is what the temperature feels like to the human body and accounts for relative humidity in addition to temperature and is in degrees Fahrenheit.

Figure 1: Downloaded TCEQ Data

This screenshot shows all the data downloaded from the TCEQ website. The relevant data includes T , which stands for temperature in Fahrenheit, RH , which is relative humidity in percent, the date, the timestamp, and the temperature, relative humidity, or dew point value. The rows after the value show calculations of Heat Index in Excel based on the formulas.

After all the heat indexes were calculated, heat index was used to identify heat events. An

Data Table 1.1: TCEQ Data on Excessive Heat Events

O^{HW} (Onset date)	D^{HW} (Duration)	E^{HW} (End date)
June 19, 2023	2 days	June 21, 2023
June 24, 2023	5 days	June 29, 2023
July 10, 2023	4 days	July 14, 2023
July 17, 2023	4 days	July 21, 2023
July 31, 2023	14 days	August 13, 2023
August 17, 2023	9 days	August 26, 2023
September 5, 2023	3 days	September 8, 2023

Next, urban and rural sites were identified for quantifying UHI effects. The sensor data was obtained from Dr. Lu Liang, which had the sensor ID, latitude and longitude coordinates for each sensor, date, dew point, relative humidity, temperature, and $PM_{2.5}$. The dependent variables used were temperature, or intensity of heat in a specific location, in degrees Fahrenheit. The percent

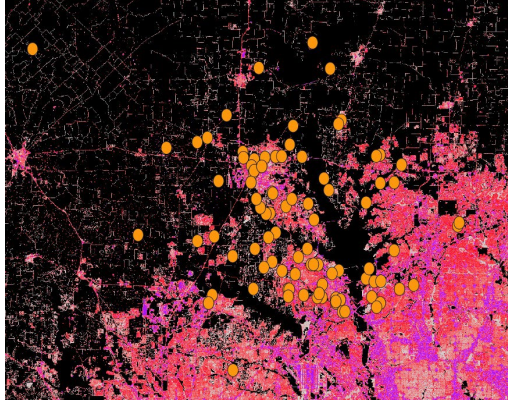
excessive heat event occurs when a “...heat index will be 105 or higher during the day and 75 or higher at night for at least a 48-hour period” (National Weather Service). Heat events can be characterized by the onset date (O^{HW}), the duration or period of consecutive heat wave days (D^{HW}), and the end date (E^{HW}).

Figure 2: Sample Combined TCEQ/NOAA Analyzed Data

Caption: Sample of combined temperature/relative humidity and sunrise/sunset data with calculated heat index (using Rothfusz’s or Steadman’s formulas) and identified heat events (1 = heat event, 0 = no heat event). Filtered from 271,786 data points

developed impervious surface map was downloaded from the National Land Cover Database. To identify urban, suburban, and rural sites, the Quantum Geographic Information System (QGIS) was used by initially uploading the impervious surface map into the program. Then, the latitude and longitude of each sensor was inputted into the program and make a buffer zone of 500 m around each sensor.

Figure 3: Sensors Plotted on NLCD Map



Caption: Zoomed in image of Denton sensors (orange) on urban imperviousness map (pink/purple = more impervious, black = less impervious).

Then, the values in the buffer zone were averaged to obtain an average imperviousness value for each sensor. The attribute table was then opened, which can be pasted into Excel. Based on the Anderson land cover classification system, which says that an area with less than 20% imperviousness is rural, 20%-50% is suburban, and 50%+ is urban, each location was classified as either urban or rural using an excel formula (Anderson, 1983). The dependent variable is the urbanization level of the location. Each sensor's classification as urban or rural was then calculated

Figure 4: Sample Classified Sensor Data

1	Timestamp	Temperature	Sensor_ID	Latitude	Longitude	Imperviousn	Classification
11	2023-01-01 00:00:00-06:00	20.7175926	113969	33.114066	-97.366056	2.09488036	Rural
12	2023-01-01 00:00:00-06:00	20.4621991	47669	33.071255	-97.027318	52.3328691	Urban
13	2023-01-01 00:00:00-06:00	21.7824074	113664	33.296769	-96.98606	25.2647305	Suburban
14	2023-01-01 00:00:00-06:00	21.0648148	113660	33.287205	-97.077165	1.82152817	Rural

Caption: Sample of the sensors classified by urbanization level based on imperviousness. Filtered from 27,996 data points.

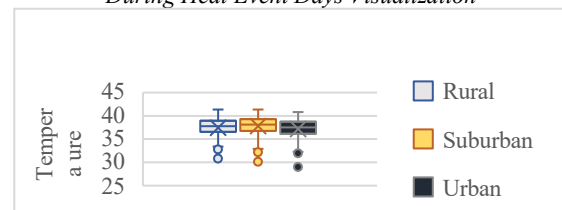
During the non-heat wave event dates, the average temperature from rural sites, suburban sites, and urban sites were calculated. Then, on the heat wave event dates, the average temperature for all three types of locations were calculated.

CONCLUSION

During heat event days, the average temperature for rural areas was 37.42 degrees Celsius, the average temperature for suburban areas was 37.77 degrees Celsius, and the average temperature for urban areas was 37.14 degrees Celsius. This contradicted the urban heat island effect because suburban areas had the hottest temperatures, rural had the second hottest temperatures, and urban areas had the lowest

temperatures. But, for the overall year, the average temperature for rural areas was 20.83 degrees Celsius, for suburban areas was 23.45 degrees Celsius, and for urban areas was 25.14 degrees Celsius. This followed the urban heat island effect because urban areas had the highest temperatures, suburban areas had the second highest temperatures, and rural areas had the highest temperatures.

Figure 5: Urbanization Level Temperature (°C) During Heat Event Days Visualization



Caption: Box plot showing the differences in temperature during heat event days.

Then, an Analysis of Variance (ANOVA) test was then performed, in order to compare the mean between three or more groups. The data was statistically significant with a p-value of 0.01585, meaning the null hypothesis was rejected (urbanization doesn't affect temperature during heat waves) and the alternate hypothesis was supported. The F-statistic was 4.17372, which is out of the region of acceptance at 95% confidence level (3.0108), further showing that the data was statistically significant. The effect size was small, with an η^2 of 0.014, which means the group explains 1.4% variance from average and a $f = 0.12$, showing that the magnitude of average difference is small. Post-hoc analysis was done with Tukey's HSD, which showed that suburban and urban areas caused the most significant differences within the data.

This proves the alternative hypothesis or that urbanization does affect temperature during heat waves. This could be explained by the fact that during heat events, Texas suburban houses use more energy than usual, which can generate more heat. In addition, many of the neighborhoods that sensors were installed in were relatively new and the green spaces are limited or young. So, the green spaces wouldn't be much help for heat mitigation. Additionally, Texas rural areas are less vegetative

than other places, especially during the summer, so the vegetation wouldn't be much help for heat mitigation as well. Finally, the urban sites had a lot of concrete, but not high rises that trap heat, which could explain why they were the coolest.

However, there were some limitations. There was sample size bias after filtration as rural sensors had a total of 130 values, suburban sensors had a total of 354 values, and urban sensors have a total of 115 values. But, since the same sensors proved the urban heat island effect for the overall year, it showed that the urban heat island effect could have been proved. Future research that includes averaging values for each heat event (instead of averaging values for all heat event dates at once). This can give more specific results and show the overall trend of the data more. Additionally, increased data size with previous summers (2013 – 2022) and places other than Denton County is needed. This can give more accurate and specific results that could be applied to places other than Denton County. Primarily, these findings show the impact of climate change in our communities. Heat events are only going to increase in prevalence with the rise of global warming, and it is important to quantify and understand the impacts of this on society and the various types of housing today.

REFERENCES

- Air Quality and monitoring*. Texas Commission on Environmental Quality. (n.d.). <https://www.tceq.texas.gov/airquality/monops>
- Anderson, J. R. (1983). A land use and land cover classification system for use with remote sensor data. USGS. <https://pubs.usgs.gov/pp/0964/report.pdf>
- California Environmental Protection Agency. (2023). *California Air Resources Board*. Inhalable Particulate Matter and Health (PM_{2.5} and PM₁₀) | California Air Resources Board. [https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health#:~:text=Particles%20are%20defined%20by%20their,diameter%20\(PM2.5\)](https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health#:~:text=Particles%20are%20defined%20by%20their,diameter%20(PM2.5)).
- Cromartie, J. (2023, October 23). *What is rural?*. USDA ERS. <https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural/#:~:text=According%20to%20the%20current%20delineation,not%20necessarily%20follow%20municipal%20boundaries>.
- Elsevier. (2015, May 29). *Recent challenges in modeling of Urban Heat Island*. Sustainable Cities and Society. <https://www.sciencedirect.com/science/article/pii/S2210670715000414>
- Fox, K., Keck, A., & Richmond, J. (2023, September 14). *NASA announces Summer 2023 hottest on record – climate change: Vital signs of the planet*. NASA. [https://climate.nasa.gov/news/3282/nasa-announces-summer-2023-hottest-on-record/#:~:text=The%20months%20of%20June%2C%20July,C\)%20warmer%20than%20the%20average](https://climate.nasa.gov/news/3282/nasa-announces-summer-2023-hottest-on-record/#:~:text=The%20months%20of%20June%2C%20July,C)%20warmer%20than%20the%20average).
- Grigorieva, E., & Lukyanets, A. (2021, June 19). *Combined effect of hot weather and outdoor air pollution on Respiratory Health: Literature Review*. MDPI. <https://www.mdpi.com/2073-4433/12/6/790/htm>
- Jia, P., Liang, L., Tan, X., Chen, J., & Chen, X. (n.d.). *Potential effects of heat waves on the population dynamics of the dengue mosquito Aedes albopictus*. PLOS Neglected Tropical Diseases. <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0007528>
- Liang, L., Daniels, J., Bailey, C., Hu, L., Phillips, R., & South, J. (2023, May 18). *Integrating low-cost sensor monitoring, satellite mapping, and Geospatial Artificial Intelligence for intra-urban air pollution predictions*. Science Direct. <https://www.sciencedirect.com/science/article/pii/S0269749123008345>
- Liang, L., & Gong, P. (2020, October 29). *Urban and air pollution: A multi-city study of long-term effects of urban landscape patterns on air quality trends*. Nature News. <https://www.nature.com/articles/s41598-020-74524-9#Sec20>.
- Multi-Resolution Land Characteristics Consortium. (2023, July 20). NLCD 2021 percent developed imperviousness (CONUS). <https://www.mrlc.gov/data/nlcd-2021-percent-developed-imperviousness-conus>
- Newcastle University. (n.d.). *Coefficient of Determination, R-squared*. Numeracy, Maths and statistics - academic skills kit. <https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/regression-and-correlation/coefficient-of-determination-r-squared.html>
- PennState. (2023). *Regression Methods*. PennState: Statistics Online Courses. <https://online.stat.psu.edu/stat501/lesson/5>
- The Heat Index Equation*. Heat index equation. (2022, May 12). https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml
- Tong, S., Prior, J., McGregor, G., Shi, X., & Kinney, P. (2021, October 25). *Urban heat: An increasing threat to Global Health*. BMJ (Clinical research ed.). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8543181/>
- United Nations. (2024). *Urbanization | Population Division*. United Nations. <https://www.un.org/development/desa/pd/content/urbanization-0#:~:text=The%20world%20is%20becoming%20increasingly,around%20two%2Dthirds%20in%202050>.
- United States Environmental Protection Agency. (2023, October 31). *Reduce urban heat island effect | US EPA*. United States Environmental Protection Agency. <https://www.epa.gov/green-infrastructure/reduce-urban-heat-island-effect>
- USA.gov. (2023, September 26). *Urban and rural*. US Census Bureau. <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html#:~:text=To%20qualify%20as%20an%20urban,population%20of%20at%20least%205%2C000>.
- US Department of Commerce, N. (2018, June 5). *Severe weather awareness - heat waves*. National Weather Service. <https://www.weather.gov/mkx/heatwaves>
- Zhao, C., Jensen, J., Weng, Q., & Weaver, R. (2018, September 7). *A geographically weighted regression analysis of the underlying factors related to the surface urban heat island phenomenon*. MDPI. <https://www.mdpi.com/2072-4292/10/9/1428>

Predicting Summer Heatwaves in the Pacific Northwest Using A Novel Integration of Multiple Machine Learning Models

Grace Zhao

ABSTRACT

Heatwaves, extended periods of extreme heat, serve as the leading cause of weather-related deaths and can cause major crop damage and detrimental economic losses. Heatwaves occur when over 2 days of a region's maximum temperature are above the 90th percentile of the daily historical average. This project investigates whether heatwaves can be predicted with above 70% accuracy for lead times greater than 1-2 weeks using a novel integration of machine learning techniques. Heatwave predictor data (precipitation, Pacific Ocean sea surface temperatures, and more) and heatwave determiner data (maximum temperature over the Pacific Northwest) from 1982 – 2023 were downloaded from public archives. The data was smoothed with a 5-day rolling mean, standardized into a daily z-score, and normalized between 0 and 1. Four neural networks were trained on the 40 years of climatological data to predict between 7 days and 42 days of heatwave predictor data, with the predictions from the best model used to test 4 classification models, which were trained on the same climatological data and compared for the highest accuracy. The Long Short-Term Memory Model neural network had a low mean squared error of 0.018 and was used to predict heatwave features, which were tested with the Support Vector Classifier that had the highest test accuracy of 96.417%. The average integrated accuracy was 74.032%. This ensemble model satisfied the project's goals of both predicting future conditions and identifying heatwaves in a novel way, giving it potential to be used in weather forecasts for timely predictions.

INTRODUCTION

1.1 Rationale

The purpose of this paper is to engineer a unique solution to predicting heatwaves in the Pacific Northwest by combining predictions from deep learning models with trained classification models. Heatwaves are currently one of the most dangerous natural hazards. Heatwaves have killed more people in the U.S. than any other type of weather (Lewis, 2021), beating hurricanes and tornadoes. During an average year from 2004 to 2018, there were 68,000 emergency department visits due to heat-related illness, 9,200

hospitalizations, and 702 deaths (Rosenthal, 2022). However, as climate change continues to fuel the increase in temperatures at a rate far higher than before, these numbers are only to increase. In 2022, there were over 1,700 deaths due to heat-related causes (CDC, 2023), a number that far surpasses that of any other natural occurrence. This is a 70% increase in heat-related deaths in just the past 4 years. Additionally, aside from human health problems, heatwaves are connected to other phenomena, such as droughts and wildfires, as well as detrimental changes associated with agriculture,

infrastructure, and biodiversity (Weirich-Benet, Elizabeth et al., 2022).

Heatwaves are an issue that affects all regions of the globe, including the Pacific Northwest (PNW) region of the United States and Canada. In 2021, the PNW experienced a major heatwave, with many locations breaking their all-time maximum temperature records by 5 °C (White, R.H. et al., 2023) and hundreds of people from British Columbia, Washington, Oregon, and California hospitalized or died due to these extreme temperatures.

In order to prevent human health worsening, economic loss, and more, a solution must be developed to predict heatwaves well in advance so that local government leaders are able to implement policies, supply resources, and help those in need. Current weather forecasting models implement dynamic models which use equations to calculate and predict upcoming temperatures. However, due to the randomness of the atmosphere, current-day atmospheric conditions are not able to be used to accurately predict temperature over 14 days later using equation-based models. This is because current atmospheric data, such as sea surface temperature, will no longer be in the atmosphere 14 days later, meaning that the data will already be too outdated for equation-based models to use and accurately create predictions from. Additionally, current ensemble forecasts are highly computationally heavy due to the large number of different models they have to run for so many different feature variations. In addition, ensemble forecasts for the Pacific Northwest are currently only able to predict up to 10 days in advance with a 70%

accuracy (White). Thus, it is crucial that new methods of prediction and forecasting are explored for heatwaves to be predicted well in advance.

In other regions of the world, such as Central Europe, researchers have been investigating the prediction of heatwaves using machine learning (ML) models such as linear regressions and random forests. They have found that for over 2 weeks of lead time, ML models have been able to compete with ECMWF forecasts, which is the forecast currently used in Europe (Weirich-Benet, Elizabeth et al., 2022). However, the usage of ML models for heatwave predictions has yet to be explored for the Pacific Northwest region. In addition, there has been no approach using an integration of deep learning models with classification models.

This paper aims to combat the rigidness of traditional dynamic and equation-based models by using an integration of machine learning models (deep learning models with classification models), which can learn from trends and may be able to reveal hidden indicators that scientists were less aware of, to predict heatwave likelihood in the Pacific Northwest weeks in advance. The engineering goals it aims to address are as follows:

- 1) Develop a neural network to be used for accurate feature prediction 7 – 42 days into the future.
 - 2) Train a classification model to be used to identify heatwaves using the neural network's feature predictions.
2. Predictor Selection
- For the training of the classification model, several predictors had to be chosen as driving factors

determining heatwaves. The following predictors were selected:

Data Table 1.1 – Data values for predictors on a specific day.

Predictor	Units; Components (if applicable)	Data Source
Precipitation	mm	NOAA
Sea Surface Temperature (SST)	5mm sea surface temperature (°C); NEP, NWP, NINO4, NINO3, NINO1+2	NOAA
Pacific North American Pattern (PNA Index)	n/a	NOAA
Integrated Water Vapor (IWV)	kg/m ²	ERA5 Reanalysis Dataset
Integrated Vapor Transport (IVT)	kg/ms	ERA5 Reanalysis Dataset
Madden-Julian Oscillation (MJO)	n/a; RMM1, RMM2	Bureau of Meteorology

Temperature was not included as a predictor, as it was used as a determiner of heatwave. However, it is to be noted that the maximum temperature (°C) was downloaded from NOAA and used to determine a day as being a heatwave day or not.

METHODOLOGY

3.1 Data Pre-Processing

The data collected all lies within the range of 1982 – 2023. Data available from public archives consistently only goes back to the year 1979. However, early weather data collected from weather station observations are inaccurate, hence why the data selection began in 1982. After the initial download, the dataset had 3.8 billion multivariate data points.

Daily data was initially downloaded for the entire world. A domain representing the Pacific Northwest was defined as being from 42° N – 53° N and 115° W – 124° W (Heeter). The data was then sliced to only incorporate that domain. A 5-day rolling mean was then calculated for each predictor using the following formula:

$$aa_{SSS} = \frac{1}{MM} \sum_{ii=0}^{nn-1} xx_{SS-ii}$$

The rolling mean was calculated in order to smoothen the data and reduce noise. This helps account for any anomalies which could affect the training of the data. Then, all the months aside from the summer months were filtered out (JJA). May and September were kept in a separate dataset in order to account for May's predictors affecting June's heatwaves and August's predictors affecting September weather patterns. Then, the data was normalized using the z-score formula. This was done in order to standardize the data to get rid of the units of measurement. The formula is as follows:

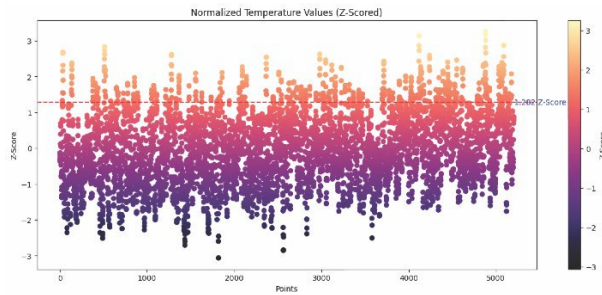
$$zz = \frac{xx - \mu\mu}{\sigma\sigma}$$

For each individual day (e.g. June 1st), the climatological average was calculated by averaging the data value of all the days of that specific date, regardless of years (e.g., June 1st 1982 – 2023 were all averaged to find the climatological mean on June 1st). Then, the climatological mean was subtracted from each individual day's value and divided by the standard deviation of that day to calculate the z-score. Finally, the data was split up into training, validation, and testing data. Separately, each value in the training, validation, and testing datasets were scaled between a scale of 0 to 1 for the neural networks, as having the weights between 0 and 1 help reduce the bias towards singular variables. There was a total of 5208 rows of data.

The above process was applied to the maximum temperature as well, with the exception of being scaled between 0 and 1. A heatwave was defined as the maximum temperature being in the 90th percentile or higher for 2 days or longer. Thus, if the z-score for the maximum temperature was above 1.282 (the representation of the 90th

percentile on a normal bell curve), and there were multiple of these high z-score values in a row, those days were classified as being part of a heatwave using a binary classification (1 for being part of a heatwave, 0 for not being part of a heatwave).

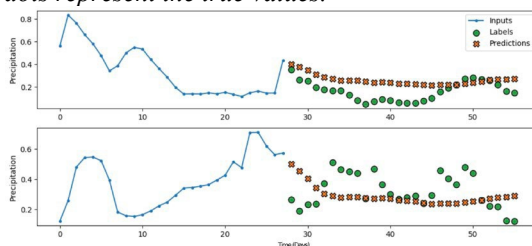
Graph 1.1 – A distribution of normalized maximum temperatures with the 1.282 z-score cutoff



3.2 Deep Learning Model Selection and Training

Four deep learning models were initially tested: the Dense Linear Model, the Convolutional Neural Network (CNN), the Long Short-Term Memory Network (LSTM), and the Auto-Regressive LSTM. After initial training, the LSTM was selected for further training and finetuning. The dense linear model only consists of a singular layer; thus, no further layers could be added to improve upon it. The auto-regressive LSTM had the highest error and was prone to overfitting due to the constant re-learning after each iteration and under-performing long-term. The LSTM was chosen over the CNN due to its ability to store past data, values, and patterns in its Cell State to use for further predictions.

Graph 1.2 – A graph showing the usage of 28 days of past data to predict 28 days of data into the future. The blue line represents the input data, the red X marks represent the values predicted by the LSTM, and the green dots represent the true values.



The LSTM was used on each of the predictors mentioned above to predict each predictor values 7, 14, 21, 28, 35, and 42 days into the future. The average mean squared error (MSE) was calculated to be 0.018, indicating that the predicted values remained relatively close to the true values. The MSE was calculated using the following formula:

$$MSE = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2$$

3.3 Classification Model Selection and Training

Four classification models were considered for this project: the logistic regression, the random forest (RF), the support vector machine (SVM), and the eXtreme Gradient Boosting Classifier (XGBoost). After initial testing, the RF and the SVM continued to be fine-tuned and trained based on having the highest initial accuracy. The best hyperparameters for each model were then sought out using the Grid-Search method. The RF model showed minimal change in the accuracy; however, the SVM accuracy increased up to 96.417%. The F1 Score was calculated to be 83.133%, the precision 83.636%, and the recall 82.635%. The formulas for each metric are as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Recall = \frac{TP}{TP + FN}$$

The best hyperparameters were determined to be a radial basis function (RBF) kernel, a gamma of 15, and a C value of 20. The SVM is able to

classify data by building a multidimensional boundary line using the RBF. The high C value indicates a stricter classification used during training and the high gamma value determines how much each point is worth. The high F1, precision, and recall scores indicate that the ratio of true positives (TP) to false positives (FP) is high, meaning that the model is not simply randomly classifying values to get a higher accuracy.

RESULTS

The results of the model integration are as follows. To integrate the LSTM and the SVM, the predictor values for each predictor and each lead time calculated from the LSTM were plugged into the trained SVM. An accuracy score was then outputted representing how correctly the model was able to predict heatwaves into the future using the calculated predictors. The accuracy was determined by how many times the model correctly classified a future time as being part of a heatwave or not.

Table 1.2 – A representation of the accuracy of the integrated models at each lead time.

Lead Time	Accuracy
7 days	73.274%
14 days	75.393%
21 days	74.277%
28 days	75.394%
35 days	73.380%
42 days	72.477%
Average	74.032%

CONCLUSION

Overall, the accuracy of the integrated model was 74.032% - a notable improvement from the current ensemble forecast accuracy of 70%. The

low MSE value of the LSTM indicates that the predictor's predicted values were close to the true values, while the SVM's high test accuracy and F-1 score, precision, and recall indicate that it has a high skill level for correctly classifying heatwaves based on features while also not simply randomly guessing. The drop in accuracy from the initial trained model to the integrated model may be attributed to the slight variation that the predicted predictor values have from their true values.

DISCUSSION

The novelty of this project lies in the integration of two different machine learning models, specifically the LSTM and the support vector machine. Additionally, the integrated model can predict up to 42 days in advance with an accuracy similar to the predictions for 7 – 14 days, indicating that it has the potential to be used for accurate long-term heatwave predictions. Although the short-term predictions are not as strong as current forecasts, the most significant finding of this research lies in the effectiveness of long-term prediction. Applications of this project can be used to accompany current weather forecasts in aiding heatwave predictions. This can then help companies and governments plan for these major weather phenomena further in advance, such as increasing manufacturing for crucial resources (air conditioners, portable fans, and more) or ensuring that everyone has adequate water access. Some current limitations of this project include the smaller number of predictors used. Additionally, the SVM may currently be overfitting, which is why the accuracy dropped when it received data that was predicted instead of true data. Thus, in the future, more predictors, such as soil moisture can be

explored, in the future. The SVM can also be finetuned further to reduce overfitting. Finally, longer lead times can be tested to determine the

boundaries of the model and how far in advance it can predict heatwaves while maintaining an accuracy above 70%

REFERENCES:

- Central for Disease Control and Prevention. (2023). Extreme Heat. CDC.
<https://www.cdc.gov/disasters/extremeheat/index.html#:~:text=Heat%2Drelated%20deaths%20and%20illnesses,Learn%20more%20about%20extreme%20heat.>
- Emerton, R., Brimicombe, C., Magnusson, L., Roberts, C., Di Napoli, C., Cloke, H. L., & Pappenberger, F. (2022). Predicting the unprecedented: forecasting the June 2021 Pacific Northwest heatwave. *Weather*, 77(8), 272-279.
- Heeter, K. J., Harley, G. L., Abatzoglou, J. T., Anchukaitis, K. J., Cook, E. R., Coulthard, B. L., ... & Homfeld, I. K. (2023). Unprecedented 21st century heat across the Pacific Northwest of North America. *Climate and Atmospheric Science*, 6(1), 5.
- Jacques-Dumas, V., Ragone, F., Borgnat, P., Abry, P., & Bouchet, F. (2022). Deep learning-based extreme heatwave forecast. *Frontiers in Climate*, 4.
- Khan, N., Shahid, S., Juneng, L., Ahmed, K., Ismail, T., Nawaz, N. (2019). Prediction of heat waves in Pakistan using quantile regression forests. *Atmospheric Research*, 221, 1-11.
<https://doi.org/10.1016/j.atmosres.2019.01.024>.
- Lewis, T. (2021). Why Extreme Heat Is So Deadly. *Scientific American*.
<https://www.scientificamerican.com/article/why-extreme-heat-is-so-deadly/>
- Puvvula, J., Abadi, A. M., Conlon, K. C., Rennie, J. J., Jones, H., & Bell, J. E. (2022). Evaluating the sensitivity of heat wave definitions among North Carolina physiographic regions. *International journal of environmental research and public health*, 19(16), 10108.
- Rosenthal, J. (2022, September 8). The Health Care Consequences of Extreme Heat Are Just the Tip of the Iceberg. Center for American Progress. Retrieved April 5, 2024, from
<https://www.americanprogress.org/article/the-health-care-consequences-of-extreme-heat-are-just-the-tip-of-the-iceberg/#:~:text=The%20inequitable%20impacts%20of%20extreme%20heat&text=According%20to%20the%20CDC%2C%20in,702%20heat%2Drelated%20deaths%20occurred.>
- Weirich-Benet, E., Pyrina, M., Jiménez-Estève, B., Fraenkel, E., Cohen, J., & Domeisen, D. I. V. (2023). Subseasonal Prediction of Central European Summer Heatwaves with Linear and Random Forest Machine Learning Models. *Artificial Intelligence for the Earth Systems*, 2(2), e220038.
<https://doi.org/10.1175/AIES-D-22-0038.1>
- White, R. H., Anderson, S., Booth, J. F., Braich, G., Draeger, C., Fei, C., ... & West, G. (2023). The unprecedented Pacific Northwest heatwave of June 2021. *Nature Communications*, 14(1), 727.
- Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). *Dive into Deep Learning*. Cambridge University Press. Retrieved from <https://D2L.ai>.

Settlement Preferences and Development of Aurelia Labiata Planula Larvae, on Natural and Anthropogenic Substrates in Current and Climate Change Temperature Scenarios

Zoe Kecskemeti

INTRODUCTION

This study aims to determine a method to mitigate jellyfish blooms, specifically in urban coastal areas. As sea temperatures rise and the coastal regions become more populated, the question arises: Will these factors result in a greater increase in jellyfish aggregations? To answer this question, the settling behavior of *Aurelia labiata* planulae was observed and measured to determine their preferences for settling on artificial or natural substrates in varying temperatures.

The *Aurelia labiata* (Figure 1) is part of the *Aurelia* genus, commonly called “moon jellyfish”. The *Aurelia* genus is within the Cnidaria phylum, class



Figure 1. Photo of *Aurelia labiata* taken at the Point Defiance Zoo and Aquarium.

Scyphozoa. Medusae, or mature jellyfish from the *Aurelia* genus, are present globally, directly impacting ecological and anthropogenic environments (Hoover & Purcell, 2008). The *Aurelia labiata* is found in the Pacific Ocean in coastal areas from Alaska to California (Gershwin, 2001; Widmer, 2005). They are a pelagic species, meaning they are

found in the upper 1,000 meters of the ocean. The temperature tolerances of *A. labiata* are estimated to be between -6 and 31 degrees Celsius (MacMullan, 2005), but they have a preferred temperature range of 2 to 26 degrees Celsius (Fletcher, 2019). Medusae from the genus *Aurelia* are mostly translucent, with a white outline. They have a dome-shaped bell as large as 40 centimeters in diameter, with usually four horseshoe-shaped reproductive organs, or gonads, in the center of the bell.

The *Aurelia labiata*’s life cycle (Figure 2) is metagenetic, meaning they reproduce sexually and asexually (Brekman et al., 2015). Sexual reproduction occurs in the medusa stage and begins when the eggs in the female’s manubrium are fertilized. These then become planulae, which are covered in tiny hairs called cilia to propel them through the water. Once released from the female medusa, the planulae develop for a few days then

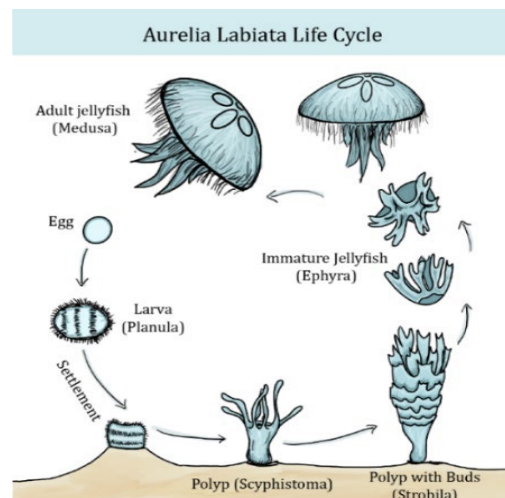


Figure 2. *Aurelia labiata* life cycle

settle on a solid substrate. Once settled, the planulae transform into scyphistomas or polyps (Collins, 2018). Planulae often settle on the underside of surfaces because those that have settled upright have shown instances of food residue decaying in their gastric cavity, causing suffocation or attracting predators, resulting in the death of the polyp. When on the underside of substrates, the food residue does not build up around the polyp (Holst & Jarms, 2006). Additionally, gravity works in favor of the polyp's tentacles, allowing them to reach further for food.

A. labiata feeds on zooplankton, including copepods, crustaceans, small fish, and other jellyfish (MacMullan, 2005). An *A. labiata* medusa catches its prey through its nematocyst-covered tentacles. The prey is injected with a toxin released from the nematocysts and becomes paralyzed. Once the prey is paralyzed, the medusa uses its oral arms to bring the prey to its mouth (Fidous, 2023). *A. labiata* are harmless to humans but, in large numbers, can negatively impact the ecosystem and the daily lives of many.

Jellyfish in genus *Aurelia* are prone to large aggregations, where their populations increase drastically in a short period. Worldwide, jellyfish aggregations, or blooms, have been increasing in both frequency and intensity (Hoover & Purcell, 2008). Jellyfish blooms have many harmful outcomes in both natural and anthropogenic settings. Zooplankton densities within aggregations in Quatermaster Harbor and Sinclair Inlet were 73% lower than in other areas (Hickey, 2022). As jellyfish populations grow, increasing competition makes planktivorous species, such as clams, whales, and fish, more vulnerable. In addition, jellyfish blooms

create disadvantages for anthropogenic activities. Fishermen worldwide have reported clogging and breaking of fishing nets due to jellyfish. Another issue is high catch mortality and damage to the gills and skin of fish due to the jellyfish's nematocysts (Bosch-Belmar et al., 2020). Moreover, aggregations have resulted in nuclear power plant shutdowns in the United States, Japan, Sweden, South Korea, Scotland, the Philippines, and Israel. To regulate the power plant's temperature, pipes connecting to the sea are used to uptake the water, so when clogged with jellyfish, the power plant is at risk of overheating (D'Agostino, 2021).

While jellyfish have thrived for millennia, the changing ocean conditions favor their populations. A prominent reason for the rise of jellyfish aggregations is anthropogenic impacts. The 43 billion tons of carbon emitted each year damages the health of the oceans through acidification. Jellyfish have a higher tolerance to ocean acidification than their predator species, allowing their population to increase. Furthermore, carbon dioxide and other greenhouse gases contribute to global warming, 90% of which is up taken by the oceans (NASA). When greenhouse gases are emitted, the infrared radiation entering Earth becomes trapped in the atmosphere and heats the planet. While global warming and cooling are natural phenomena, human activities have accelerated the process, so Earth's natural mitigation system can no longer return the planet to a stable state. Alongside climate change, humans are disturbing marine ecosystems through the development and urbanization of coastal zones. In January of 2023, worldwide, 2.15 billion people lived in near-coastal areas and 898 million in low-

elevation coastal zones. These numbers could rise to 2.9 billion and 1.2 billion by 2100 (Reimann et al., 2023). As more people occupy these areas, the plastic, paper, and other anthropogenic waste in the oceans increases. Marine life is forced to adapt to these conditions to survive and are harmed significantly in the process. Again, in this situation, jellyfish are at an advantage; their predators' populations decrease. Medusae can live in altered habitats, while other species may be unable to withstand and adapt to the changing conditions. This allows the jellyfish to thrive. The floating waste gives the jellyfish planulae, or larvae, a place to settle that was not present before so they can then reproduce more efficiently and with greater success (Holst & Jarms, 2006).

Climate change and human development are occurring simultaneously, and studies have shown that individually, they both play a role in causing jellyfish to bloom. When asking why jellyfish populations suddenly spike, looking at the early stages of their life cycle, planulae settlement is one way to approach the question. If planulae become increasingly successful in their settlement, they will have a higher chance of successful reproduction. There have been studies on the artificial substrate preferences of planulae (Holst & Jarms, 2006; Hoover & Purcell, 2008), as well as how temperature impacts their settlement (Webster & Lucas, 2012), but not the combined effect. To better understand the projected jellyfish populations, it is essential to look at both.

This study hypothesizes that the substrates will not impact the planulae's viability. Jellyfish aggregations typically occur in coastal areas, so planulae are exposed to various substrates. It also

hypothesizes that temperature and substrate collectively impact the settlement of the planula, and the planulae will settle more on artificial substrates as temperatures increase. This is because of the effects temperatures have on different substrates. For instance, artificial substrates tend to be designed to withstand a broader range of temperatures with minimal effects on the substrate. In contrast, natural substrates are specifically suited for their native environments and will see more significant differences.

METHODS

Pilot Study

Before completing this study, a pilot study was conducted using barnacle shell, rock, PET, concrete, and galvanized steel. 12 well plates were used; 2 designated for control, and each of the 5 substrates was present on 2 plates. The substrates were tested individually at 13°C, a known temperature that planulae will develop, to ensure that none would have external effects on the planulae and influence the results. Once the wells were populated with planulae and the substrate (not including the control), the planulae were left to settle for one week.

At the end of the week, the non-viable planulae in each well were counted and taken from the mean number of planulae in each well calculated during set-up. Non-viable planulae are those which would be unable to contribute to a population. These may be shriveled, significantly asymmetrical, or not beginning to transition into the next life cycle stage. Based on this mean, the number of viable planulae was placed in **Table 1**, which shows a sample of the data tables used for this study.

Substrates	Barnacle Shell	Rock	PET	Concrete	Galvanized Steel	Control
Well 1	59	59	57	61	2	61
Well 2	59	60	56	61	8	59
Well 3	61	59	61	60	11	60
Well 4	61	60	55	59	10	60
Well 5	60	59	61	59	4	60
Well 6	61	58	61	60	3	60
Well 7	59	58	57	60	12	61
Well 8	59	60	58	61	3	61
Well 9	61	59	59	60	4	59
Well 10	61	58	60	61	3	61
Well 11	60	61	59	59	2	60
Well 12	61	57	59	61	4	61

Table 1. Number of Viable Planulae with Five Substrates and a Control

Preconditioning of Substrates

The substrates used for this study included galvanized steel, polyethylene terephthalate (PET), concrete, barnacle shells, and rocks collected from Puget Sound beaches. Enough of each of these substrates was obtained to create 54 approximately 10 mm x 10 mm x 5 mm rectangular samples. The natural substrates were not in these exact dimensions but were kept relatively the same size. The substrates placed in the wells were picked intentionally to be similar sizes. All materials were kept in natural seawater for a week to develop a biogenic film from the bacteria in the natural seawater. Planulae settle quicker and more successfully when the substrates are coated with a biofilm.

Extraction of Planulae

This experiment used a stock culture of *Aurelia labiata* (**Figure 3**) from the Point Defiance Zoo and Aquarium. These medusae were sourced from Vashon Island, Washington. Using a plastic transfer pipette, 40 samples of planulae were

extracted from the oral arms of randomly selected female *Aurelia labiata* and transferred into a 1-L plastic beaker filled with 25 mL of filtered seawater. The collected samples were poured through a 1000-micron mesh strainer into another 1-L plastic beaker and then through a 250-micron mesh strainer into a 400-mL glass beaker. This way, the medusae's mucus and oral arm parts were filtered out, and the samples were left with planulae only. The glass beaker was filled with 150 mL of filtered seawater. An aeration tube was placed in the glass beaker to create a current using vigorous bubbling. After placing the aeration tube in the glass beaker, the planulae were left alone for 5 minutes to disperse (Figure 4).

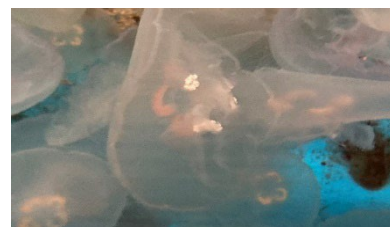


Figure 3. Planulae producing female *Aurelia labiata* from the stock culture at the Point Defiance Zoo and Aquarium. The white cluster on the underside of the medusa is the planulae.

Populating the Wells

A SCILOGEX P20 micropipette set at 20 μ L was used to collect the planulae from the glass beaker. A BIOLOGIX 6-well sterile cell culture plate was placed under a Zeiss STMI 508 microscope. 3 to 4-20 μ L samples were taken from the glass beaker and placed into one of the six wells, attempting to get between 55 and 65 planulae per well. This process was repeated for the remaining 5 wells, and the well plate was labeled 15°C, trial 1. This was then repeated for trials 2 and 3 and again for all three trials for well plates for the 20°C trial and 25°C trial.

Experiment Set Up

The 4 substrates, PET, shell, rock, and concrete, were placed randomly in each of the planulae-populated wells of the 9 plates. The wells were then filled 5/6 full of filtered seawater, about 14 mL, and the lid was placed on the plates. The plates were then grouped into their respective temperatures.

3 Lucky Reptile Herp Nursery II temperature incubators were used for this experiment. 1 USB LASCAR temperature data logger was placed inside each of these incubators.

The temperature of the incubators was set so the LASCAR data loggers read 15°C, 20°C, and 25°C. The average temperature of each incubator was recorded for the duration of the experiment.

Once the incubators were calibrated to their assigned temperatures, the plates were placed into the incubator, matching the temperature of their label. The incubators were closed, and the planulae were left to settle for two weeks.

Data Collection

Once the two-week settlement period ended, the plates were removed one at a time from the incubator. One plate was moved under the microscope, and the viable planulae, those still swimming or beginning to develop without deformities, were counted. Settled planulae were also counted, which were no longer moving and had begun to develop tentacles. The settled planulae were recorded in tables similar to Table 2 for the substrate they were settled on. The planulae settled on the water tension and began to grow tentacles near the water's surface in the well plate. These planulae were counted as control and recorded in the respective (15°C, 20°C, or 25°C) data table.

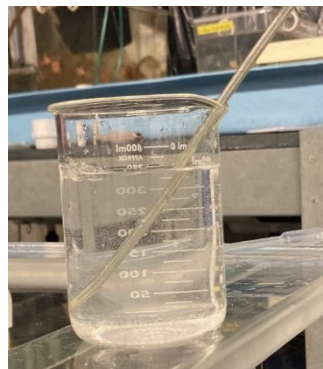


Figure 4. An aeration tube is placed into a 400 mL glass beaker filled with diluted planula water. The aeration tube is causing the water to bubble vigorously creating a current which circulates the water and disperses the planula.

Table 2. Number of Planulae Settled on Substrates and Water Tension at 15 Degrees Celsius on the First Well Plate

<i>Substrates</i>	<i>Barnacle Shell</i>	<i>Rock</i>	<i>PET</i>	<i>Concrete</i>	<i>Well Plate</i>	<i>Water Tension</i>	<i>Not Settled</i>
Well 1	0	3	6	0	4	5	41
Well 2	0	2	1	2	5	6	45
Well 3	0	0	23	0	1	3	32
Well 4	1	4	18	1	0	2	36
Well 5	1	1	10	0	6	8	42
Well 6	2	3	12	0	0	0	43

RESULTS

Pilot Study All data was analyzed using RStudio. The pilot study was intended to ensure that the substrates would not affect the viability of the planulae. The data from **Table 1** was imported into RStudio, and a box plot was created to show the difference between the viable planulae for each substrate.

Additionally, a two-sample t-test, which is used to compare the means of two separate groups, was used to compare the results of each substrate to the results of the control. The null hypothesis for the t-test is that there is no significant difference between the two groups.

Substrate	P-value
<i>Galvanized Steel</i>	$2.853e^{-15}$
<i>PET</i>	0.0181
<i>Concrete</i>	0.7999
<i>Rock</i>	0.0048
<i>Shell</i>	0.8127

Table 3 The p-values of the t-test compare the means of the viable planulae on each substrate to the means of viable planulae in the control.

Table 3 lists the significance of each t-test, and the null hypothesis was accepted for concrete and shell, meaning the viable planulae with these two substrates did not differ from the control. The

remaining three substrates, rock, PET, and galvanized steel, showed significant differences from the control. However, in Figure 6, the difference in means for PET, rock, and the control

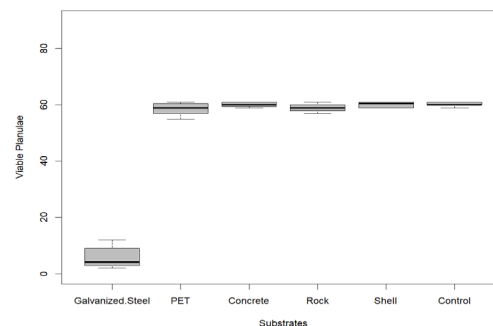


Figure 5. Viable Planulae on Five Substrates and a Control. Bars show the minimum, maximum, mean, first quartile, and third quartile of the viable planulae in each of the 12-well samples.

was minimal, not exceeding 2 planulae, indicating little difference between the samples. Therefore, although the t-test showed a significant difference from the control, the substrates were kept in the final study. The significant p-values, indicating a difference in the mean values, were likely due to small sample sizes, low variation within each sample, and human error in counting the planulae. However, galvanized steel showed greater significance, with a p-value below 0.0001 and a substantial difference in the means. The mean values of the control and galvanized steel differed by almost 55 planulae. Figure 6 demonstrates that the viable planulae in the galvanized steel plates deviate from the trend shown by the remaining substrates and the control. This difference in viability between the control and galvanized steel samples resulted in the exclusion of galvanized steel in the multiple substrate study. Based on these results, if the steel was included with the rest of the substrates, there would be low viability in the wells, and no usable data could be collected from the experiment.

Multiple Substrate Study

Data from the multiple substrate study was imported into RStudio, and temperature and substrate were created as factor variables used to categorize data. To analyze this data, 4 generalized linear models were created, as these models allow building a linear relationship with non-linear data. A linear relationship shows how well variables correlate, in this case, how well settlement would correlate to temperature or substrate. The generalized linear models used in this analysis had three parts: response variable, predictor variable (or

variables), and were of the family Poisson, the type of model used for count data. All models had the number of settled planulae as the response variable, the dependent variable, and the factors temperature and substrate as predictor variables. The first model examined the correlation of temperature and substrate as codependent variables and settlement. The second model had temperature and substrate as independent variables as predictors for settlement. The third model considered only temperature, and the fourth used substrate as a predictor variable.

4 diagnostic plots were created to measure if the data meets the expectations of the model being used to analyze it. The first plot, Residuals vs. Fitted, shows a horizontal or approximately horizontal line to indicate a linear pattern among the residuals. The second plot is the Q-Q residuals plot, which checks for normal distribution. The data is normally distributed if the points are generally on a straight, diagonal line. The third plot is the Scale-Location plot, and this plot checks for homoscedasticity or equal variance in the data. If the line across the plot is approximately horizontal, the variance in the data is relatively equal (Zach, 2021). The fourth plot is the Residuals versus Factor levels plot, which also checks for homoscedasticity. The variance is equal if the points on this plot are scattered evenly from the horizontal line. The plot for this model suggests that the assumptions have been met. However, the Scale-Location plot suggests there may be some unequal variance. Although this is inconsistent in the Residuals versus Factor levels plot, that also checks for homoscedasticity.

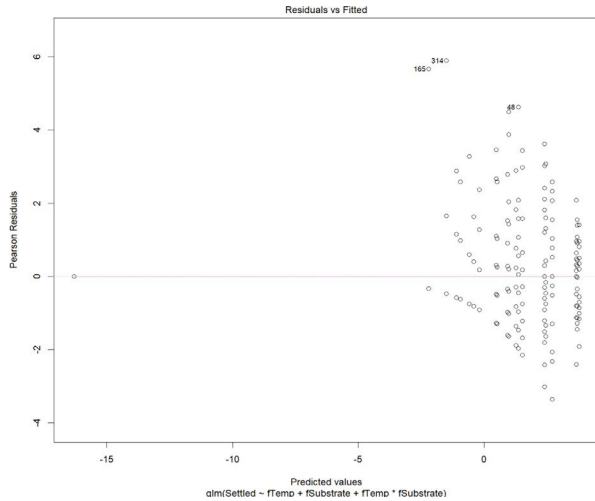


Figure 7. Residuals vs Fitted plot. The line across the center is horizontal which means that the data fits the assumption of a linear pattern.

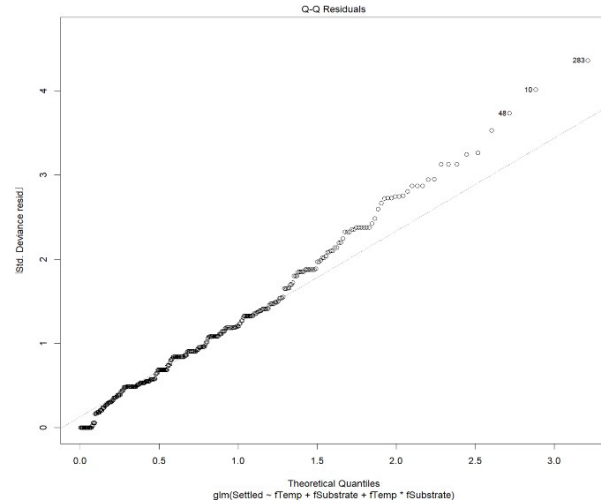


Figure 8. Q-Q Residuals plot shows that the data stays relatively close to the dashed diagonal line and appears to be slightly right skewed.

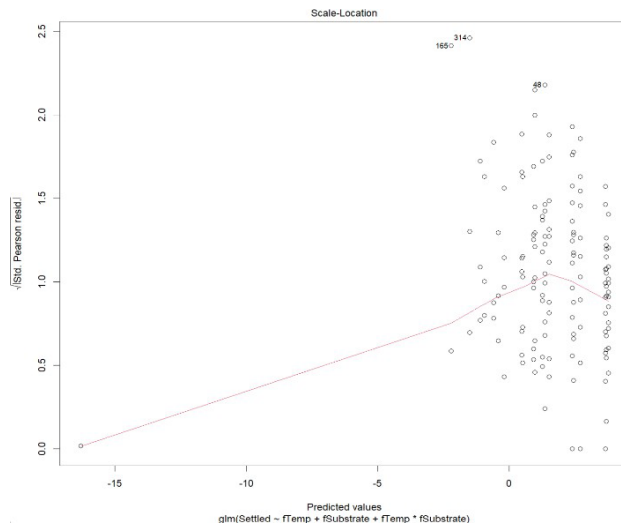


Figure 9. The Scale-Location plot shows somewhat deviating variance, and points clustered mostly to the right, indicating that there may be some unequal variance.

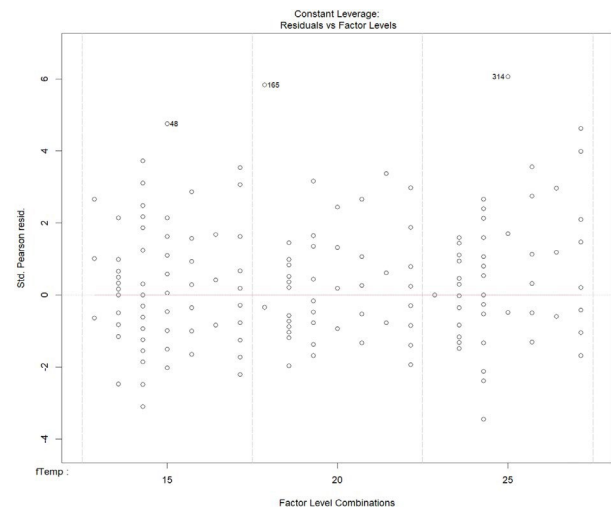


Figure 10. The Residuals versus Factor Levels plot shows equal mostly variance of the points from the line $r=0$.

Three methods were used to determine the model best fitting for the data. The first was an Akaike Information Criterion, or an AIC: $2KK - 2\ln(LL)$, where K is the number of independent variables used and L is the log-likelihood estimate (Bevans, 2020). Log-likelihood is a value that shows the probability that the model can predict the dependent variable. The lower the AIC is for a

model, the greater its accuracy. The resulting values of the AIC test on the four models were 1575.131 for model 1, 1679.135 for model 2, 7776.5 for model 3, and 1675.946 for model 4. The lowest AIC value, 1575.5, was for the first model, looking at the codependent relationship between temperature and settlement.

Model	Degrees of Freedom	Residual Deviance	P-Value
Null	377	6853.3	
Temperature	2	6852.4	0.6665
Substrate	6	743.1	<0.0001
Temperature x Substrate	12	615.1	<0.0001

Table 4. Results of the ANOVA test on the most significant model, the interaction between substrate and temperature

Table 4 shows the results of the ANOVA, and temperature alone was not significant in the settlement of planulae. The substrate temperature interaction had the lowest residual deviance, which is why this model was the best fit for the data.

Secondly, an ANOVA was used for nested regression models, which tests the significance of more complex models using the Chi-squared test. A low residual deviance indicates that a model predicts the response variable using the given predictor variables well. This ANOVA test checks if making the model more complex by adding another predictor variable or creating an interaction between predictor variables significantly lower residual deviance.

Third, a pseudo-R-squared value using McFadden's formula was calculated for each model. R-squared is also used to test for the fit of a model, and it is a value that ranges between 0 and 1. An R-squared value close to one means the model fits the

data and that the independent variables account for the variation, or how the data is spread from the mean, caused by the dependent variable.

McFadden's formula calculates an R-squared value for generalized linear models that do not give an R-squared value, such as Poisson regression. This value is calculated using the following formula:

$1 - \frac{\text{deviance of full model}}{\text{deviance of null model}}$, which, based on the results in **Table 4**, would be $1 - \frac{615.1}{6853.3}$. The calculated value is approximately 0.91, showing that the interaction between temperature and substrate mostly accounts for the variance in settlement.

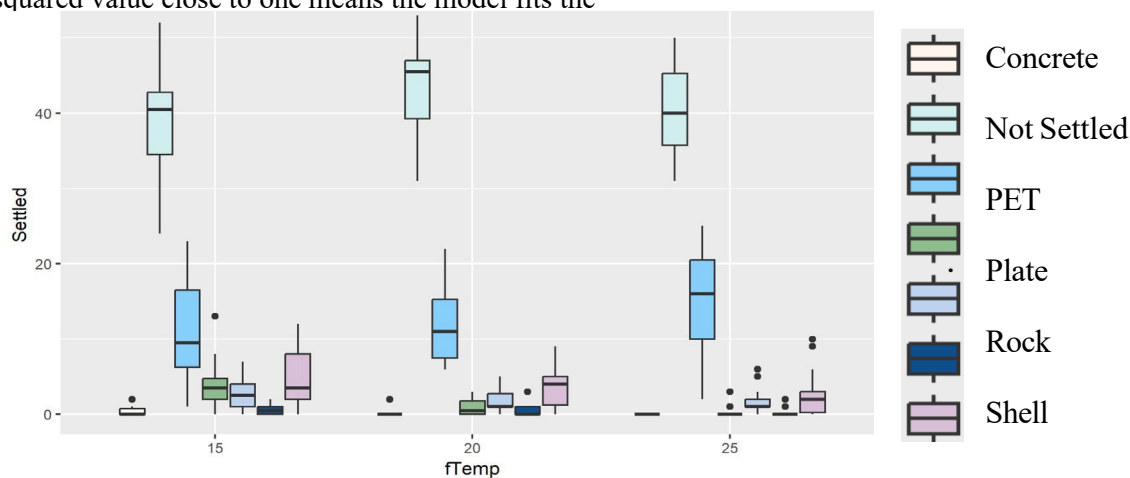


Figure 6. Planulae Settlement on Substrates in 15°C, 20°C, and 25°C. The colors of the box and whiskers plot represent the substrate, the height shows the number settled on that substrate, and each of the three sections is a temperature.

Figure 11 illustrates the planulae’s settlement in the 3 different temperatures. The “Not Settled” category has the largest number of planulae because the planulae that settle quickly will feed on those not yet settled. “Not Settled” represents the number of planulae not on substrates, the well plate, or the surface tension of the water. No planulae were left swimming in the wells after the 2-week settlement period, so the planulae on substrates were counted and added. That number was subtracted from the

original number of planulae in the well to determine the number that was not settled. From looking at the graph, additional trends in the data include that as temperatures increased, the mean number of planulae settled on PET increased, while settlement on the well plate, shell, and concrete decreased. To determine whether these increases and decreases were significant, t-tests between the 15°C and 20°C, the 20°C and 25°C, and the 15°C and 25°C groups of each substrate were completed.

Substrates	15°C vs. 20°C	20°C vs. 25°C	15°C vs 25°C
Concrete	0.298	0.331	0.055
PET	0.739	0.094	0.078
Rock	0.212	0.925	0.164
Not Settled	0.041	0.119	0.602
Well Plate	0.001	0.037	0.0001
Shell	0.576	0.361	0.149
Water Tension	0.365	0.419	0.058

Table 5. P-values of t-tests done on different temperature groups of substrates.

The only significant differences between settlement in different temperatures were found on the well plate, where decreases in settlement were consistently significant, and the increase in the number of planulae not settled between 15°C and 20°C, seen bolded in Table 5.

DISCUSSION

The pilot study demonstrated that planulae have low viability when exposed to a controlled environment containing galvanized steel (p-value $2.853e^{-15}$). This refutes the first hypothesis used in the study that the substrate would not affect the planula’s viability. Zinc is a heavy metal that negatively impacts aquatic invertebrates (Jeong et al., 2023) and is used in galvanizing (coating) steel. Galvanized steel is used in coastal development, for

instance, on dock pillars and latches, and does not impair the growth of algae, mollusks, and marine bacteria. Using this material for infrastructure rather than plastics or wood may decrease jellyfish populations. Seeing the effect of galvanized steel on the vitality of planulae, the settlement of planulae on structures such as pillars would decrease. While galvanized steel offers potential benefits in mitigating jellyfish aggregation, its production process, including mining and galvanization, can negatively affect the environment.

The hypothesis of the multiple substrate study was partially correct, where temperature and substrate codependently were the best predictors of settlement. Although, there was no significant change in the preference for artificial substrates. The

study remained consistent with previous studies where plastic, in this case PET, was the most desirable substrate. PET consistently had the most planulae settled in the three temperatures. While there was an increase in the means of the settled planulae, it was not significant. The PET was floating on the surface of the well plate while the remaining substrates were on the bottom. This may have influenced where the planulae chose to settle. However, the recorded number of planulae on the surface tension of the water was less than those on PET. This indicates that the PET had a greater influence on the settlement than its placement. In addition, PET is found in the ocean on the bottom of docks, plastic water bottles, and certain fibers, which float on the water's surface and create an ideal place to settle for the planula larvae.

The well plate, also made of plastic, had comparatively low settlement to PET. A possible explanation for this is the preconditioning of the substrates. The substrates were coated in a bacterial (biogenic) film after being soaked in natural seawater for a week, while the well plate was not. The lack of biogenic film may have discouraged the planulae from settling, and why their settlement decreased as water temperatures increased. Bacteria populations grow larger in warmer temperatures, so there would be more biogenic film in warmer temperatures. The greater bacterial film on the substrates is likely why the settlement on the well plate decreased significantly with increasing temperatures.

The natural substrates and concrete were the least desired for settlement. The concrete used in this

study was made partially of natural substrates, which may have affected its desirability. The non-uniform surfaces of natural substrates may explain why they are not preferred for settlement. Flat surfaces, such as PET, provide more areas where the planulae can settle, while on natural substrates, it is more challenging due to the grooves and ridges. Using natural substrates like shell and rock, or even a combination of natural and artificial substrates, would, to an extent, reduce the anthropogenic impact on the ecosystem, having it be closer to mimicking the natural environment and therefore causing less strain on the species in coastal areas. In addition, these could mitigate jellyfish blooms without the negative ecological impacts of galvanized steel.

Future Research

While this study may help to provide insight into ways to reduce jellyfish blooms and their impact, it does not give a definite solution to returning jellyfish aggregations to their normal frequencies and intensities. Additional research would need to be done on settlement and other life cycle stages, exploring different substrates and combinations of different climate change or anthropogenic-related causes of these aggregations. For instance, how will hypoxia increase the competition between jellyfish and surrounding species due to increased algal blooms? As jellyfish have been seen to have a higher tolerance to changing ocean conditions, are there any parts in their lifecycle that are seen to be the largest dictator in survival? If so, what can be done to reduce rapid population increases?

REFERENCES

- Bevans, R. (2020, March 26). *Akaike Information Criterion | When & How to Use It*. Scribbr. <https://www.scribbr.com/statistics/akaike-information-criterion/>
- Bosch-Belmar, M., Milisenda, G., Basso, L., Doyle, T. K., Leone, A., & Piraino, S. (2020). Jellyfish Impacts on Marine Aquaculture and Fisheries. *Reviews in Fisheries Science & Aquaculture*, 29(2), 1–18. <https://doi.org/10.1080/23308249.2020.1806201>
- Brekthman, V., Malik, A., Haas, B., Sher, N., & Lotan, T. (2015). Transcriptome profiling of the dynamic life cycle of the scypohozoan jellyfish *Aurelia aurita*. *BMC Genomics*, 16(1), 74. <https://doi.org/10.1186/s12864-015-1320-z>
- Brewer, R. H. (1978). Larval Settlement Behavior in the Jellyfish *Aurelia aurita* (Linnaeus) (Scyphozoa: Semaestomeae). *Estuaries*, 1(2), 120. <https://doi.org/10.2307/1351601>
- Brewer, R. H. (1984). THE INFLUENCE OF THE ORIENTATION, ROUGHNESS, AND WETTABILITY OF SOLID SURFACES ON THE BEHAVIOR AND ATTACHMENT OF PLANULAE OF *CYANEA* (CNIDARIA: SCYPHOZOA). *The Biological Bulletin*, 166(1), 11–21. <https://doi.org/10.2307/1541426>
- Collins, A. (2018, April). Jellyfish and Comb Jellies | Smithsonian Ocean. *Ocean.si.edu*. https://ocean.si.edu/ocean-life/invertebrates/jellyfish-and-comb-jellies#section_16508
- D’Agostino, S. (2021, October 28). *Jellyfish attack nuclear power plants. Again and again*. The Bulletin; Bulletin of the Atomic Scientists. <https://thebulletin.org/2021/10/jellyfish-attack-nuclear-power-plant-again/>
- Firdous, F. (2023, May 2). Moon Jellyfish: Habitat, Diet & Other Facts. *Ocean Fauna*. <https://oceanfauna.com/moon-jellyfish/>
- Fletcher, M. (2019). *Aurelia*. *Animal Diversity Web*; University of Michigan. <https://animaldiversity.org/accounts/Aurelia/>
- Gershwin, L. (2001). Systematics and Biogeography of the Jellyfish *Aurelia labiata* (Cnidaria: Scyphozoa). *The Biological Bulletin*, 201(1), 104–119. <https://doi.org/10.2307/1543531>
- Hassan, M. (2023, May 31). ANOVA (Analysis of variance) - Formulas, Types, and Examples. *Research Method*. <https://researchmethod.net/anova/>
- Hickey, H. (2022, March 3). Moon jellies appear to be gobbling up zooplankton in Puget Sound. *UW News*. <https://www.washington.edu/news/2022/03/03/moon-jellies-appear-to-be-gobbling-up-zooplankton-in-puget-sound/>
- Holst, S., & Jarms, G. (2006). Substrate choice and settlement preferences of planula larvae of five Scyphozoa (Cnidaria) from German Bight, North Sea. *Marine Biology*, 151(3), 863–871. <https://doi.org/10.1007/s00227-006-0530-y>
- Hoover, R. A., & Purcell, J. E. (2008). Substrate preferences of scyphozoan *Aurelia labiata* polyps among common dock-building materials. *Hydrobiologia*, 616(1), 259–267. <https://doi.org/10.1007/s10750-008-9595-6>
- Jeong, H., Byeon, E., Kim, D.-H., Maszczyk, P., & Lee, J.-S. (2023). Heavy metals and metalloid in aquatic invertebrates: A review of single/mixed forms, combination with other pollutants, and environmental factors. *Marine Pollution Bulletin*, 191, 114959. <https://doi.org/10.1016/j.marpolbul.2023.114959>
- MacMullan, C. (2005). *Aurelia labiata* (Moon jellyfish). *Animal Diversity Web*; University of Michigan. https://animaldiversity.org/accounts/Aurelia_labiata/
- NASA Global Climate Change. (2022). Ocean Heat Content | NASA Global Climate Change. Climate Change: Vital Signs of the Planet. <https://climate.nasa.gov/vital-signs/ocean-warming/>
- Reimann, L., Vafeidis, A. T., & Honsel, L. E. (2023). Population development as a driver of coastal risk: Current trends and future pathways. *Cambridge Prisms: Coastal Futures*, 1. <https://doi.org/10.1017/cft.2023.3>
- Tomaru, A., Sasaki, R., Miyahara, H., Okino, A., Ogawa, N., & Hamasaki, K. (2014). Settlement of Planulae of the Moon Jellyfish *Aurelia aurita* onto Hydrophilic Polycarbonate Plates Modified by Atmospheric Plasma Treatment. *PLoS ONE*, 9(1), e85569. <https://doi.org/10.1371/journal.pone.0085569>
- Webster, C. N., & Lucas, C. H. (2012). The effects of food and temperature on settlement of *Aurelia aurita* planula larvae and subsequent somatic growth. *Journal of Experimental Marine Biology and Ecology*, pp. 436–437, 50–55. <https://doi.org/10.1016/j.jembe.2012.08.014>
- Widmer, C. L. (2005). Effects of temperature on growth of north-east Pacific moon jellyfish ephyrae, *Aurelia labiata* (Cnidaria: Scyphozoa). *Journal of the Marine Biological Association of the United Kingdom*, 85(3), 569–573. <https://doi.org/10.1017/s0025315405011495>

AWARDS

Envirothon

Regionals, State and Nationals March 2024 – August 2024

King County (KingCo) Envirothon Woodinville, Washington (March 2024)

1st Place

Envirothon Team Pigeon

Aashray Chegu Vijay
Saharsh Bhargava
Aaryan Dhingra

Tanmay Sarvepalli
Madeleine Vanderheiden
Yuquing Wang

Envirothon Team Sparrow

Aanchal Batwara
Yuzhu Guo
Joliz Karukayil
Saketh Madiraju

Eeshan Misra
Alexander Peev
Sanskriti Sinha

Washington State Envirothon Camp Wooten Pomeroy, Washington (May 2024)

1st Place

Envirothon Team Pigeon

Aashray Chegu Vijay
Saharsh Bhargava
Aaryan Dhingra

Tanmay Sarvepalli
Madeleine Vanderheiden
Yuquing Wang

National Envirothon (August 2024)

Tesla STEM: “Outstanding Oral Presentation” award

Exploravision National Competition Honorable Mention

“Oncotype DXDL”

“Syntax: Creation of Semisynthetic Paclitaxel to Mitigate Decreasing Biological Availability of T. brevifolia and Drug Costs”

Future Business Leaders of America (FBLA)

State and Nationals April 2024 – July 2025

Future Business Leaders of America (FBLA) State Spokane, Washington (April 2024)

Accounting II

14th Place Sarah Wen

Advertising

1st Place Samyuktha Krishnan

Banking & Financial Systems

2nd Place Varun Hari, Sriram Marasanapalle-Kalle

Broadcast Journalism

4th Place Anshu Aggarwal, Rishi Dilip, Saathvik Somujayabalan

5th Place Elaine Jiang, Madeline Tsai

8th Place Rhea Singh, Sophie Zhao

Business Calculations

2nd Place Pranav Kannepalli

9th Place Divya Ramu

Business Communication

4th Place Samyuktha Krishnan

Business Ethics

10th Place Ayaan Kumar, Siddhesh Nikam, Kanishk Yadav

Business Law

3rd Place Madeline Tsai

Business Management

1st Place Nitya Bhartiya, Saahiti Mamidipaka

Business Plan

1st Place Varsha Mantravadi, Shriya Tiku

6th Place Alekhya Nidadavolu, Harshita Raja, Sahaana Sridhar

Coding & Programming

5th Place Nimai Belur, Vasudha Narayanan, Aarushi Verma

11th Place Aakarsh Balla, Sarthak Lodha, Sacchin Saravanan

Community Service Group Project

11th Place

Computer Game & Simulation Programming

6th Place Jai Bindlish, Aashrith Pisupati, Pragyan Singh

Computer Problem Solving

3rd Place Alexander Nyczkowski

13th Place Elaine Jiang

Cyber Security

1st Place Siddhesh Nikam

3rd Place Alexander Nyczkowski

18th Place Sarah Wen

Data Analysis

2nd Place Aashi Juneja, Roxana Stanescu

Digital Animation

12th Place Pearly Agarwal, Anshu Aggarwal, Pawan Grewal

Digital Video Production

7th Place Alexa Rorrer, Sarah Wen

12th Place Shreeyaa Muhunthan, Keerthi Rajesh, Pallavi Thiruveedu

Economics

15th Place Diya Balaji

Electronic Career Portfolio

5th Place Keerthi Rajesh

Financial Statement Analysis

3rd Place Ekaansh Agrawal, Neil Parikh, Jaskirat Singh

8th Place Varshini Hari, Layla Kaim

Future Business Educator

1st Place Kaavya Trivedi

Impromptu Speaking

1st Place Kaavya Trivedi

Insurance & Risk Management

10th Place Theo Petchdenlarp

11th Place Aashi Juneja

International Business

2nd Place Richa Badi, Diya Balaji

Introduction to Business Concepts

6th Place Sophie Zhao

Introduction to Business Presentation

1st Place Varshini Hari, Layla Kaim

11th Place Ekaansh Agrawal, Neil Parikh, Jaskirat Singh

Introduction to Business Procedures

15th Place Veda Hiremath

Introduction to Financial Math

6th Place Neil Parikh

7th Place Max Xie

8th Place Jai Bindlish

Introduction to Information Technology

3rd Place Hem Akarapu

12th Place Aarush Kapoor

Introduction to Programming

5th Place Kaushal Dabbiru, Safwan Hasan, Tejas Panja

7th Place Yohaana Ardeni, Akshath Dommaraju, Sathkrith Segu

Introduction to Social Media Strategy

5th Place Aarushi Bhatia, Rachel Cai, Madison Desilva

Local Chapter Annual Business Report

4th Place Group Entry

Marketing

3rd Place Anusha Pandey, Divya Ramu

Mobile Application Development

8th Place Achintya Agrawal, Pranav Kannepalli, Abhinav Vallabhaneni

12th Place Likitha Nanduri, Divya Ramu, Sanskriti Sinha

Network Design

8th Place Eshaal Omar, Shruthika Saravanakumar, Milana Veera

Networking Infrastructures

19th Place Ayaan Kumar

Organizational Leadership

9th Place Megan Halim

Personal Finance

11th Place Shikha Singh

Public Policy & Advocacy

20th Place Divij Aswinkumar

Public Service Announcement

4th Place Saahiti Mamidipaka, Harshita Raja

Sales Presentation

1st Place Pavitra Dave, Kaavya Trivedi

Securities & Investments

7th Place Aakarsh Balla

Social Media Strategies

9th Place Ava Arvind, Sarah Youn

Sports & Entertainment Management

5th Place Divij Aswinkumar, Nanda Junu, Pranay Nookala

UX Design

3rd Place Grace Zhao

11th Place Sahaana Sridhar

Visual Design

8th Place Prisha Agarwal, Anjana Punniamoorthy

10th Place Aparna Iyer, Saathvik Somujayabalan

Website Coding & Development

2nd Place Aakarsh Balla, Sacchin Saravanan, Abhinav Vallabhaneni

Website Design

1st Place Rohan Baliga, Rajvir Mishra

**Future Business Leaders of American (FBLA) Nationals
Orlando, Florida (June 2024)**

Cybersecurity

7th Place Alex Nyczkowski

Public Service Announcement

8th Place Harshita Raja and Sahaana Sridhar

Electronic Career Portfolio

3rd Place Keerthi Rajesh

Intro to Financial Math

4th Place Max Xie

Sales Presentation

2nd Place Pavitra Dave and Kaavya Trivedi

Health Occupations Students of America (HOSA)

State and Nationals March - June 2024

Health Occupations Students of America (HOSA) State,

Spokane, Washington March 2024

Clinical Specialty

1st Place Inaya Chowdhury

2nd Place Joliz Karukayil

7th Place Rudra Malatpure

Creative Problem Solving

6th Place Saachi Sharma, Anushka Adhikary

Extemporaneous Writing- Health Policy

1st Place Akchara Mukunthu
3rd Place Shreeyaa Muhunthan

Family Medicine Physician

3rd Place Pavitra Dave

Health Career Display

7th Place Anisha Upasani, Shipra Prajapati
9th Place Veda Hiremath, Preetika Kaur

Health Career Photography

3rd Place Basudha Ghose
10th Place Sophia Rodriguez

Health Informatics

4th Place Akchara Mukunthu

Healthy Living

6th Place Basudha Ghose

Human Growth & Development

1st Place Samyuktha Krishnan
5th Place Inaya Chowdhury

Job Seeking Skills

5th Place Abhinav Nadupalli

Medical Innovation

6th Place Essa Ali, Ajeet Venkatesh
8th Place Kenisha Jaiswal, Moomal Siddiqui, Aparna Iyer, Talah Albadawi
10th Place Shruthika Saravanakumar, Milana Veera

Medical Reading

3rd Place Abhinav Nadupalli

Mental Health Promotion

1st Place Shriya Krishan, Aanchal Batwara
5th Place Deva Mitra Junu, Prisha Agarwal

Nutrition

1st Place Samyuktha Krishnan
10th Place Defne Dingiloglu

Pathophysiology

1st Place Berke Yilmaz

Pharmacology

1st Place Samyuktha Krishnan
6th Place Rajiv Arvind

Physical Therapy Health Professions Events

8th Place Inaya Chowdhury

Prepared Speaking

7th Place Harshita Raja
8th Place Aahana Mittal

Public Health

1st Place Rajiv Arvind, Diya Modi
9th Place Samyuktha Krishnan, Defne Dingiloglu, Anugraha Pillai

Public Service Announcement

10th Place Winston Zhou, Anirudh Narasmihan, Siddharth Mulupuru, Hans Rehet, Shravan Mahesh Kumar

Researched Persuasive Writing and Speaking

2nd Place Priyanka Kudallur

**Health Occupations Students of America (HOSA) Nationals
Dallas, Texas (June 2024)**

Extemporaneous Writing-Health Policy

4th Place Shreeyaa Muhunthan
8th Place Akchara Mukunthu

Health Career Photography

10th Place Basudha Ghose

Researched Persuasive Writing and Speaking

1st Place Priyanka Kudallur

Emergency Preparedness Events

10th Place Shriya Krishnan, Aanchal Batwara

Barbara James Service Award:

Bronze Level Aahana Mittal

**Holocaust Center for Humanities
Art Contest**

1st Place Likitha Nanduri, *“In tribute to Eva Tannenbaum Cummins' indomitable spirit, this piece is a representation of her remarkable journey. Emerging from shattered glass, reminiscent of a bullet hole, is the silhouette of triumph—a portrayal of Eva transcending tragedy to shine as a celebrated theater actress. Beyond the fractured barrier lies a desolate landscape, haunted by echoes of loss— Kristallnacht's remnants, the destruction of thousands of synagogues, her father's death, and leaving her homeland. Amidst this barrenness, delicate rose petals intertwine, symbolizing Eva's reverence for her past and her dedication to honoring and respecting Holocaust victims. Though Eva passed in 2022, through her advocacy, the echoes of the past are heard, encouraging remembrance and resilience for generations to come.”*

LWPTSA REFLECTIONS – “I am Hopeful Because...”

*WSPTA Reflections finalist

Outstanding Interpretation

Dance:	Anugraha Pillai*	“Redemption”
Music Composition:	Diya Modi*	“Dandelions”
Visual Arts:	Vivian Li*	“Confidence from Within”

Award of Excellence

Visual Arts:	Tanvi Patil	“Lifeline”
--------------	-------------	------------

Award of Merit

Music Composition:	Diya Modi	“Dandelions”
--------------------	-----------	--------------

Math Competitions

Math Prize for Girls (November 2023)

Silver Medal Reese Long

American Invitational Mathematics Examination (AIME) Qualifiers

Ekaansh Agrawal
Aakarsh Balla
Jai Bindlish
Pranav Kannepalli
Reese Long

Rui Meng
Alex Peev
Sathkrith Segu
Sarah Wen
Max Xie

MUSIC

WIAA/Washington Music Educators Associations State Solo and Ensemble Contest (February 2024)

2nd Place Flute/Piccolo, Yilin Zhang

Northlake Solo & Ensemble Festival (March 2024)

1st Place Small Woodwind Ensemble *Fantastic Flutes*, Manami Chiba, Yilin Zhang
1st Place Flute/Piccolo, Manami Chiba
1st Place Flute/Piccolo, Yilin Zhang

National Economics Challenge State (April 2024)

3rd Place Kaavya Trivedi, Adhip Ghosh, Alekhya Nidadavolu
4th Place Vivaan Bansal, Aditya Nautiyal, Richa Badi, Diya Balaji

National History Day (Regionals & State, January-June 2024)

National History Day Regionals (January 2024)

1st Place Samantha Novitski & Saachi Sharma
 "Discovering Taxol: How One Tree Revolutionized Modern Cancer Treatment"
1st Place Cindy Yang
 "The Codebreakers of Bletchley Park: A Turning Point for WWII"
2nd Place Anson Chen
 "Embracing a Free World: Eisenhower, the 1952 Republican National Convention and the Birth of Bipartisan Internationalism"
3rd Place Samuel Yang
 "Embracing a Free World: Eisenhower, the 1952 Republican National Convention and the Birth of Bipartisan Internationalism"

National History Day State Kirkland, Washington (April 2024)

*Advanced to Nationals

1st Place *Cindy Yang
 Lyle L. Nelson Research Award
 Awarded \$150
 "The Codebreakers of Bletchley Park: A Turning Point for WWII"
2nd Place *Samuel Yang
 "USAF System 621B's Role in Modern GPS"

5th Place Anson Chen
"Embracing a Free World: Eisenhower, the 1952 Republican National Convention and the Birth of Bipartisan Internationalism"

National Merit Semi-Finalists

September 2023

Chegu Aashray	Shauryan Deb	Chris Ma	Anusha Pandey
Anshu Aggarwal	Aparna Iyer	Varsha Mantravadi	Divya Ramu
Saraswathy Amjith	John Karaca	Vasudha Narayanan	Saathvik Somujayabalan
Nimai Belur	Rithvik Koppolu	Aylin Ozdemir	Colin Xie

Science and Engineering Fairs

CSRSEF, WSSEF and ISEF March – May 2024

Central Sound Regional Science and Engineering Fair (CSRSEF)

Regionals (March 2024) *Advanced to Finalist Judging on Day 2

Behavioral and Social Sciences

2nd Place Nitya Bhartiya*
"Color Me Surprised Identifying and Evaluating Logo Color Trends' Impacts on Business Brand Awareness"

4th Place Kamden Applebee
"Interplay of Individual Personality and Language Background on Speech Accommodation"

Biological Sciences and Biomedical Engineering

1st Place Saachi Sharma and Nikki Taleghani*
"Optimizing upper extremity positioning to increase respiratory function in multiple sclerosis patients with COPD"

Biomedical and Health Sciences

2nd Place Pavitra Dave*
"Investigating the Effect of Varying Vitamin D Solutions on Human Cancer T-cell Proliferation and IL-2 Secretion"

3rd Place Keshav Iyer
"Understanding the Relationship Between the Gut Microbes and Leukemia Growth"

4th Place Ajeet Venkatesh
"MIP Based Monoclonal Antibodies for Effective Cancer Immunotherapy"

Biochemistry

1st Place Aryaman Rtunjay
"ALS-SynAegis A Molecular Dynamics Study on TDP-43 Aggregation to Prevent Amyotrophic Lateral Sclerosis Onset"

Computational Biology and Bioinformatics

4th Place Berke Yilmaz
"Machine Learning Framework for Early Detection and Progression Tracking in Sporadic ALS"

Earth and Environmental Sciences

1st Place Anson Chen
"Forecasting Domoic Acid Levels from Harmful Algal Blooms along the Pacific Northwest Coast"

4th Place Grace Zhao
"Predicting Summer Heatwaves in the Pacific Northwest Using Machine Learning"

Engineering Technology

- 1st Place Nikhil Kadam*
 “Development of a Non-Invasive Biosensor Device for Early Detection of Periodontal Diseases and Caries”
- 4th Place Alex Takagi Stewart
 “Minimization of Thermoacoustic Combustion Instability in Exploration-Class Rocket Engines”

Material Science

- 1st Place Harshita Raja*
 “O-Patch Developing an Omega-3 Microneedle Patch for Women with Postpartum Depression”
- 4th Place Eamon McGarty
 “Creating a Graphene Based Fluorescent Fingerprint Powder”

Additional Awards

ISEF Finalist

Aryaman Rtunjay
Anson Chen

Yale Science and Engineering Special Category Award

Aryaman Rtunjay

Ricoh Sustainability Development Special Category Award

Sanjana Katti

American Institute of Aeronautics and Astronautics Special Category Award

Alex Takagi Stewart

Washington State Science and Engineering Fair (WSSEF) State Bremerton HS, Bremerton, Washington (April 2024)

Behavioral and Social Sciences

- 1st Place:
- Edbert Chenjp
 “Are Asian Families Putting Too Much Pressure On Their Children”
- Aahana Mittal
 “A Quantitative Analysis Exploring the Correlation Between Anxiety Dynamics and Academic Performance Among Individuals of Asian Descent”
- Alexander Zhang
 “Childhood Impact on Cross-Race Bias in Witness Scenarios”
- Priyanka Kudullar
 “Competitive Learning & Cognitive Resilience Analyzing Peers’ Influence on Problem-Solving Skills Using Electroencephalography Signals”
- Kamden Applebee
 “Interplay of Individual Personality and Language Background on Speech Accommodation”
- Inaya Chowdhury
 “Testing the Effectiveness of Progressive Muscle Relaxation on the Cognitive Performance of Adolescents with Somatic Anxiety”
- Sahaana Sridhar
 “Voice of Anxiety Utilizing Acoustic Biomarkers and Machine Learning for Enhanced Anxiety Diagnosis”

Saranya Pammi

“Carnatic Musopathy for Inducing Quality Sleep”

Anugraha Pillai

“Can Changes in Core-Body Temperature Be Predicted Using Machine Learning

2nd Place:

Nitya Bhartiya

“Color Me Surprised Identifying and Evaluating Logo Color Trends’ Impacts on Business Brand Awareness”

Alexa Rorrer *“Reel Equality A Quantitative Content Analysis of the Twenty Highest-Grossing Films of All Time”*

Marilyn Zhao

“Is it Actually That Dang Phone The Effect of Social Media Usage on Emotional Intensity in Response To Art”

3rd Place

Elizabeth James

“Stress Reduction with the Use of an App that Promotes Mindful Breathing”

Biomedical and Health Sciences

1st Place:

Pavitra Dave

“Supplementing Immunity Investigating the Effect of Varying Vitamin D Solutions on Human Cancer T-cell Proliferation and IL-2 Secretion”

Ajeet Venkatesh, *“MIP Based Monoclonal Antibodies for Effective Cancer Immunotherapy”*

Keshav Iyer, *“Understanding the Relationship Between the Gut Microbes and Leukemia Growth”*

Marta Jayantilal, *“Viability of Saliva for Hashimoto’s Thyroiditis Diagnosis”*

Bhanu Atmakuri, *“Optimizing 2000m Rowing Performance An Experimental Study Assessing Ergometer Performance in Relation to Pacing Strategy”*

Sophia Holmes, *“Why Does My Dad Need His Coffee A Study of Caffeine Withdrawal and Age”*

Riya Singh, *“Investigating the Effect of Flavonoid Supplementation on Autism Spectrum Disorder Symptoms”*

2nd Place

Arjun Rao

“The Effect of Sublethal Concentrations of Antibiotics on Antibiotic Resistance Development in E-coli”

Computational Biology and Bioinformatics

1st Place:

Shruthika Balasubramanian, *“A Comprehensive Analysis of the Correlation Between Demographic Factors and Glaucoma Prevalence and Risk”*

Defne Dingiloglu, *“Elucidating Streptomyces Caviscabies’ Biosynthetic Product Potential Through Genome-Mining Analysis”*

Teyj Krishnan, *“A BCI-Based EEG-Controlled Prosthetic Limb”*

2nd Place:

Arina Pastukhova

“It’s In Your Yogurt A Novel Treatment for Methicillin-Resistant S. Aureus”

Mehuli Nath

“Predicting Parkinson’s Disease Severity Using Voice Analysis”

Suvan Pal and Adhith Vutukri

“A Convolutional Neural Network Approach for Brain Tumor Detection in MRI Scans”

Earth and Environmental Sciences

1st Place:

Zoe Kecskemeti

“Settlement Preferences of Aurelia Labiata Planula Larvae in Climate Change Scenarios”

Grace Zhao
"Predicting Summer Heatwaves in the Pacific Northwest Using Machine Learning"
 Sanjana Katti
"Effects of PFOA and Ocean Warming on Coral Symbiosis"
 Eva Jain
"Impact of Heat Events on Urban Heat Island Effects"
 Regan Hendrickson
"Investigating the Effects of Biodegradable Plastic Breakdown on Mycorrhizal Fungi Growth"

2nd Place Samyuktha Krishnan
"MarineAid Mercury Assessment and Resilience Investigation in Aquatic Domains"

Engineering Technology

1st Place: Alex Takagi Stewart
"Minimization of Thermoacoustic Combustion Instability in Exploration-Class Rocket Engines"
 Nikhil Kadam
"Development of a Non-Invasive Biosensor Device for Early Detection of Periodontal Diseases"

2nd Place Shaan Gada
"SWAFER Safety Steering Wheel Advancements for Enhanced Roadway Safety"

Material Science

1st Place: Harshita Raja
"O-Patch Developing an Omega-3 Microneedle Patch for Women with Postpartum Depression"
 Adhip Ghosh
"Investigation of the Elastocaloric Effect in Shape Memory Alloys"

2nd Place Shuhui Yang
"The Osmoprotective Effects of Trehalose and Sucrose on Mung Bean Growth Under NaCl Stress"

Microbiology

1st Place: Saahiti Mamidipaka
"Investigating the Success of Microbial Treatment for Gluten Digestion"
 Megan Halim
"Dynamics of Female Gut Colonization by Uropathogenic E Coli"
 Eda Keskin
"The Antibacterial Effect of Ohi'a Lehua Blossom Honey on Escherichia Coli"

Additional Awards

American Psychological Association Award

Edbert Chenjp
 Aahana Mittal

Ohio Wesleyan University Outstanding Achievements in Science \$80,000 Scholarship

Aahana Mittal
 Priyanka Kudullar
 Kamden Applebee
 Ajeet Venkatesh

Keshav Iyer
Marta Jayantilal

Association for Women Geoscientists Award

Grace Zhao
Sanjana Katti
Eva Jain

Regeneron International Science and Engineering Fair (ISEF) Finalist

Ajeet Venkatesh
Keshav Iyer

**International Science Engineering Fair (ISEF)
Los Angeles, California (May 2024)**

Biochemistry

4th Place Aryaman Rtunjay
 “ALS-SynAegis A Molecular Dynamics Study on TDP-43 Aggregation to Prevent Amyotrophic Lateral Sclerosis Onset”

Earth and Environmental Sciences

2nd Place Anson Chen
 “Forecasting Domoic Acid Levels from Harmful Algal Blooms Along the Pacific Northwest Coast”

Computational Biology and Bioinformatics

4th Place Ajeet Venkatesh
 “T-STM Spatio Temporal Modeling of the Tumor-Microenvironment for in silico Testing of Immunomodulatory Drug Combinations for Enhanced Triple Negative Breast Cancer Immunotherapy”

Regeneron Science Talent Search Finalists 2023

Finalists

Saraswathy Amjith
“A Novel Integrated Machine Learning Approach Utilizing Radar and Satellite Imagery for Selective Logging Remote Sensing Detection and Accompanying AI-Logging Map-Generating Webtool”

Regeneron Semi-Finalists

Saraswathy Amjith
“A Novel Integrated Machine Learning Approach Utilizing Radar and Satellite Imagery for Selective Logging Remote Sensing Detection and Accompanying AI-Logging Map-Generating Webtool”

Ashwin Kaliyaperumal,
“A Quantum Optimization Algorithm To Efficiently Route Public Transportation in Cities”

**Science Olympiad
November 2023 – July 2024**

**Camas High School Science Olympiad Invitational
Camas, Washington (November 2023)**

Team Awards

1 st Place	Team Nikola
5 th Place	Team Curie
12 th Place	Team Richter
17 th Place	Team Franklin
21 st Place	Team Carver

Individual Awards

Air Trajectory

4th Place Shrey Goel, Mehuli, Nath (Team Curie)

Anatomy and Physiology

2nd Place Shuhui Yang, Angela Wang (Team Nikola)

4th Place Adhrith Vutukuri, Himay Marada (Team Richter)

Chemistry Lab

3rd Place Aaryan Dhingra & Rajit Joshi (Team Nikola)

5th Place Parikshith Venkataraghavan, Farzad Hasan (Team Curie)

Codebusters

2nd Place Olivia Wang, Divij Aswinkumar, Rohan Viswanathan (Team Nikola)

Brenda Li, Katya Rukhlinskaia, Keerthi Rajesh (Team Curie)

Detector Building

1st Place Farzad Hasan, Parikshith Venkataraghavan (Team Curie)

Disease Detectives

4th Place Grace Zhao, Ashwin Kaliyaperumal (Team Nikola)

Dynamic Planet

1st Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)

3rd Place Sathkrith Segu, Shipra Prajapati (Team Richter)

Ecology

2nd Place Olivia Wang, Divij Aswinkumar (Team Nikola)

3rd Place Shrey Goel, Madeleine Vanderheiden (Team Curie)

Experimental Design

5th Place Sree Sarvepalli, Divij Aswinkumar, Sachita Ghosh (Team Nikola)

Fermi Questions

1st Place Grace Zhao, Sree Sarvepalli (Team Nikola)

3rd Place Tyler Chen, Mehuli Nath (Team Curie)

Flight

1st Place Varun Hari, Himay Marada (Team Richter)

4th Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)

Forensics

4th Place Olivia Wang, Aaryan Dhingra (Team Nikola)

Forestry

4th Place Olivia Wang, Neha Madanagopal (Team Nikola)

Fossils

3rd Place Grace Zhao, Sachita Ghosh (Team Nikola)

Geo Mapping

2nd Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)

Microbe Mission

2nd Place Shuhui Yang, Angela Wang (Team Nikola)

Quantum Quandries

3rd Place Ashwin Kaliyaperumal, Rajit Joshi (Team Nikola)

Sustainable Energy

1st Place Sree Sarvepalli, Ashwin Kaliyaperumal (Team Nikola)

3rd Place Shrey Goel, Ansh Bayappu (Team Curie)

Robot Tour

1st Place Saharsh Bhargava, Rohan Viswanathan (Team Nikola)

Write-it Do-it

1st Place Grace Zhao, Shuhui Yang (Team Nikola)

Boyceville Science Olympiad Invitational (December 2023)**Team Awards**

4th Place Team Nikola

Individual Awards**Air Trajectory**

11th Place Sree Sarvepalli, Sriram Marasanapalle-Kalle (Team Curie)

Anatomy and Physiology

8th Place Shuhui Yang, Angela Wang (Team Nikola)

Astronomy

14th Place Aashray Chegu Vijay, Rajit Joshi (Team Nikola)

Code Analysis

13th Place Rajit Joshi, Divij Aswinkumar (Team Nikola)

Codebusters

14th Place Olivia Wang, Divij Aswinkumar, Rohan Viswanathan (Team Nikola)

Detector Building

8th Place Farzad Hasan, Parikshith Venkataraghavan (Team Curie)

10th Place Aaryan Dhingra, Divij Aswinkumar (Team Nikola)

Disease Detectives

19th Place Grace Zhao, Ashwin Kaliyaperumal (Team Nikola)

Experimental Design

5th Place Sree Sarvepalli, Divij Aswinkumar, Sachita Ghosh (Team Nikola)

Flight

13th Place Ansh Bayappu, Himay Marada (Team Curie)

15th Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)

Forensics

3rd Place Olivia Wang, Aaryan Dhingra (Team Nikola)

Forestry

11th Place Olivia Wang, Neha Madanagopal (Team Nikola)

Fossils

12th Place Grace Zhao, Sachita Ghosh (Team Nikola)

Geocaching

8th Place Grace Zhao, Ashwin Kaliyaperumal (Team Nikola)

Geo Mapping

19th Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)

Precision Medicine

2nd Place Shuhui Yang, Sree Sarvepalli (Team Nikola)

Robot Tour

1st Place Saharsh Bhargava, Rohan Viswanathan (Team Nikola)

12th Place Farzad Hasan, Ansh Bayappu (Team Curie)

Scrambler

14th Place Novyanna Bajracharya, Sia Gupta (Team Franklin)

Tower

10th Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)

Write-it Do-it

10th Place Grace Zhao, Shuhui Yang (Team Nikola)

Tesla STEM Science Olympiad Invitational (TI-24)

Executive Directors - Sree Sarvepalli, Ashwin Kaliyaperumal

Tournament Directors - Olivia Wang, Neha Madanagopal, Aashray Chegu Vijay, Grace Zhao, Saharsh, Bhargava, Shuhui Yang, Sriram Marasanapalle-Kalle

**Raisbeck Aviation HS Science Olympiad Invitational
Tukwila, WA (January 2024)****Team Awards**

1st Place Team Nikola
3rd Place Team Curie
9th Place Team Franklin
13th Place Team Richter
19th Place Team Carver

Individual Awards**Air Trajectory**

1st Place Sree Sarvepalli, Sriram Marasanapalle-Kalle (Team Curie)

Anatomy and Physiology

2nd Place Brenda Li, Himay Marada (Team Curie)
3rd Place Shuhui Yang, Angela Wang (Team Nikola)
5th Place Keerthi Rajesh (Team Richter)

Astronomy

2nd Place Shipra Prajapati, Suvan Pal (Team Curie)
3rd Place Aashray Chegu Vijay, Rajit Joshi (Team Nikola)

Chemistry Lab

2nd Place Aaryan Dhingra, Rajit Joshi (Team Nikola)

Codebusters

2nd Place Olivia Wang, Divij Aswinkumar, Rohan Viswanathan (Team Nikola)

Detector Building

1st Place Farzad Hasan, Parikshith Venkataraghavan (Team Curie)
2nd Place Divij Aswinkumar, Aaryan Dhingra (Team Nikola)

Disease Detectives

5th Place Grace Zhao, Ashwin Kaliyaperumal (Team Nikola)

Dynamic Planet

2nd Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)
5th Place Shipra Prajapati, Madeleine Vanderheiden (Team Curie)

Ecology

1st Place Brenda Li, Madeleine Vanderheiden (Team Curie)
2nd Place Olivia Wang, Divij Aswinkumar (Team Nikola)

Fermi Questions

2nd Place Grace Zhao, Sree Sarvepalli (Team Nikola)
4th Place Sia Gupta, Alexander Peev (Team Franklin)
5th Place Mehuli Nath, Tyler Chen (Team Curie)

Flight

1st Place Iris Dey, Varun Hari (Team Richter)
2nd Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)

Forensics

1st Place Olivia Wang, Aaryan Dhingra (Team Nikola)
5th Place Shreyansi Swain, Emily Li (Team Carver)

Forestry

1st Place Olivia Wang, Neha Madanagopal (Team Nikola)

Fossils

3rd Place Grace Zhao, Sachita Ghosh (Team Nikola)
4th Place Aakash Ghosh & Tejas Panja (Team Carver)

Geo Mapping

2nd Place Aashray Chegu Vijay & Saharsh Bhargava (Team Nikola)
3rd Place Shivank Mohanty & Shipra Prajapati (Team Curie)

Microbe Mission

2nd Place Shuhui Yang, Angela Wang (Team Nikola)
4th Place Mehuli Nath & Ansh Bayappu (Team Curie)

Optics

1st Place Suvan Pal & Yash Gupta (Team Curie)
2nd Place Ashwin Kaliyaperumal & Rajit Joshi (Team Nikola)
4th Place Sia Gupta & Emily Li (Team Franklin)

Quantum Quandaries

2nd Place Ashwin Kaliyaperumal & Rajit Joshi (Team Nikola)
4th Place Yash Gupta & Parikshith Venkataraghavan (Team Curie)

Robot Tour

1st Place Saharsh Bhargava & Rohan Visnawathan (Team Nikola)
4th Place Farzad Hasan & Ansh Bayappu (Team Curie)

Scrambler

1st Place Farzad Hasan & Parikshith Venkataraghavan (Team Curie)
5th Place Shreya Santhosh, (Team Richter)

Sustainable Energy

3rd Place Shrey Goel & Ansh Bayappu (Team Curie)

**Golden Gate (GSSO) Invitational, University of California Berkley
Berkley, CA (February 2024)**

Team Awards

11th Place Team Nikola

Individual Awards

Air Trajectory

8th Place Sree Sarvepalli, Sriram Marasanapalle-Kalle

Cybersecurity

11th Place Ashwin Kaliyaperumal, Rajit Joshi

Dynamic Planet

9th Place Saharsh Bhargava, Aashray Chegu Vijay

Experimental Design

13th Place Sree Sarvepalli, Sachita Ghosh, Divij Aswinkumar

Forensics

10th Place Olivia Wang, Aaryan Dhingra

Forestry

13th Place Olivia Wang, Neha Madanagopal

Fossils

15th Place Grace Zhao, Sachita Ghosh

Geo Mapping

8th Place Aashray Chegu Vijay & Saharsh Bhargava

Microbe Mission

12th Place Shuhui Yang, Angela Wang

Robot Tour

3rd Place Saharsh Bhargava & Rohan Visnawathan

Tower

11th Place Sriram Marasanapalle-Kalle, Neha Madanagopal

Write-it Do-it

10th Place Shuhui Yang, Angela Wang

Science Olympiad, Washington North West Regional, Bothell High School, Bothell, WA (March 2024)

1st Place Team Nikola
4th Place Team Curie
12th Place Team Wegener
17th Place Team Franklin
19th Place Team Carver

Individual Awards

Air Trajectory

2nd Place Shrey Goel, Ansh Bayappu (Team Curie)
3rd Place Sree Sarvepalli & Sriram Marasanapalle-Kalle (Team Nikola)

Anatomy and Physiology

2nd Place Shuhui Yang, Angela Wang (Team Nikola)
4th Place Brenda Li, Himay Marada (Team Curie)

Astronomy

2nd Place Suvan Pal & Shipra Prajapati (Team Curie)
3rd Place Aashray Chegu Vijay & Rajit Joshi (Team Nikola)

Chemistry Lab

3rd Place Aaryan Dhingra & Rajit Joshi (Team Nikola)
5th Place Parikshith Venkataraghavan, Tyler Chen (Team Curie)

Codebusters

1st Place Olivia Wang, Divij Aswinkumar, Rohan Viswanathan (Team Nikola)
3rd Place Brenda Li, Katya Rukhlinskaia, Saranya Voruganti (Team Curie)
5th Place Keerthi Rajesh, Shreya Santhosh, Shivank Mohanty (Team Wegener)

Detector Building

2nd Place Farzad Hasan, Parikshith Venkataraghavan (Team Curie)
3rd Place Divij Aswinkumar & Aaryan Dhingra (Team Nikola)

Disease Detectives

1st Place Grace Zhao, Ashwin Kaliyaperumal (Team Nikola)
5th Place Katya Rukhlinskaia & Adhrith Vutukuri (Team Curie)

Dynamic Planet

4th Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)
5th Place Shipra Prajapati & Madeleine Vanderheiden (Team Curie)

Ecology

1st Place Olivia Wang, Divij Aswinkumar (Team Nikola)
2nd Place Brenda Li & Madeleine Vanderheiden (Team Curie)

Experimental Design

2nd Place Sree Sarvepalli, Divij Aswinkumar, Sachita Ghosh (Team Nikola)
4th Place Madeleine Vanderheiden, Himay Marada, Suvan Pal (Team Curie)

Fermi Questions

2nd Place Tyler Chen, Mehuli Nath (Team Curie)
6th Place Grace Zhao, Sree Sarvepalli (Team Nikola)

Flight

3rd Place Ansh Bayappu, Himay Marada (Team Curie)
4th Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)
6th Place Iris Dey, Varun Hari (Team Wegener)

Forensics

1st Place Olivia Wang, Aaryan Dhingra (Team Nikola)

Forestry

1st Place Olivia Wang, Neha Madanagopal (Team Nikola)

Fossils

1st Place Grace Zhao, Sachita Ghosh (Team Nikola)

Geo Mapping

2nd Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)

Microbe Mission

5th Place Shuhui Yang, Angela Wang (Team Nikola)

Optics

3rd Place Suvan Pal, Yash Gupta (Team Curie)
5th Place Ashwin Kaliyaperumal, Rajit Joshi (Team Nikola)

Quantum Quandries

5th Place Ashwin Kaliyaperumal, Rajit Joshi (Team Nikola)

Robot Tour

1st Place Saharsh Bhargava, Rohan Viswanathan (Team Nikola)

Scrambler

3rd Place Rohan Viswanathan, Sachita Ghosh (Team Nikola)
4th Place Farzad Hasan, Parikshith Venkataraghavan (Team Curie)
6th Place Sia Gupta, Tanisha Karthikeyan (Team Franklin)

Sustainable Energy

2nd Place Shrey Goel, Ansh Bayappu (Team Curie) **Robot Tour**
4th Place Sree Sarvepalli, Ashwin Kaliyaperumal (Team Nikola)

Tower

1st Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)

Write-it Do-it

1st Place Grace Zhao, Shuhui Yang (Team Nikola)

**Science Olympiad State, Central WA University,
Ellensburg, WA (April 2024)**

Team Awards

1st Place Team Nikola (only high school team from Washington State to qualify for the National tournament).
5th Place Team Curie (the highest placing second team at the State tournament).

Individual Awards**Air Trajectory**

3rd Place Sree Sarvepalli & Sriram Marasanapalle-Kalle (Team Nikola)
4th Place Shrey Goel, Adhith Vutukuri (Team Curie)

Anatomy and Physiology

2nd Place Shuhui Yang, Angela Wang (Team Nikola)
5th Place Brenda Li, Himay Marada (Team Curie)

Astronomy

3rd Place Aashray Chegu Vijay & Rajit Joshi (Team Nikola)
5th Place Suvan Pal & Shipra Prajapati (Team Curie)

Codebusters

1st Place Olivia Wang, Divij Aswinkumar, Rohan Viswanathan (Team Nikola)
4th Place Brenda Li, Katya Rukhlinskaia, Saranya Voruganti (Team Curie)

Disease Detectives

2nd Place Grace Zhao, Ashwin Kaliyaperumal (Team Nikola)
4th Place Katya Rukhlinskaia & Adhith Vutukuri (Team Curie)

Dynamic Planet

2nd Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)

Ecology

1st Place Olivia Wang, Divij Aswinkumar (Team Nikola)
4th Place Brenda Li & Madeleine Vanderheiden (Team Curie)

Experimental Design

1st Place Sree Sarvepalli, Divij Aswinkumar, Sachita Ghosh (Team Nikola)

Fermi Questions

3rd Place Grace Zhao, Sree Sarvepalli (Team Nikola)
5th Place Tyler Chen, Mehuli Nath (Team Curie)

Flight

2nd Place Sriram Marasanapalle-Kalle, Neha Madanagopal (Team Nikola)

5th Place Ansh Bayappu, Himay Marada (Team Curie)

Forensics

2nd Place Olivia Wang, Aaryan Dhingra (Team Nikola)

Forestry

1st Place Olivia Wang, Neha Madanagopal (Team Nikola)

Fossils

2nd Place Grace Zhao, Sachita Ghosh (Team Nikola)

Geo Mapping

2nd Place Saharsh Bhargava, Aashray Chegu Vijay (Team Nikola)

Microbe Mission

1st Place Shuhui Yang, Angela Wang (Team Nikola)

Optics

1st Place Ashwin Kaliyaperumal, Rajit Joshi (Team Nikola)

Quantum Quandries

2nd Place Ashwin Kaliyaperumal, Rajit Joshi (Team Nikola)

Robot Tour

1st Place Saharsh Bhargava, Rohan Viswanathan (Team Nikola)

Sustainable Energy

4th Place Shrey Goel, Ansh Bayappu (Team Curie)

5th Place Sree Sarvepalli & Ashwin Kaliyaperumal (Team Nikola)

Write-it Do-it

3rd Place Grace Zhao, Shuhui Yang (Team Nikola)

4th Place Mehuli Nath & Saranya Voruganti (Team Curie)

**Science Olympiad Nationals, Michigan State University
Lansing, Michigan (May 2024)**

Team Awards

Team Nikola placed 34th out of 60 teams. Top 20 finishers

Individual Awards**Robot Tour**

1st Place Saharsh Bhargava & Rohan Viswanathan

Tower

7th Place Sriram Marasanapalle-Kalle & Neha Madanagopal

Ecology

13th Place Olivia Wang & Sree Sarvepalli

Disease Detectives

15th Place Grace Zhao and Ashwin Kaliyaperumal

Dynamic Planet

16th Place Saharsh Bhargava & Aashray Chegu Vijay

Geologic Mapping

17th Place Saharsh Bhargava & Aashray Chegu Vijay

Solar Power

18th Place Suvan Pal & Shrey Goel

Anatomy & Physiology

20th Place Shuhui Yang & Angela Wang

5th Annual STEM Expo Live Event (June 2024)**Startup Track*****Fashthetics*** Ayaan Kumar, Talah Albadawiearned*Best Overall Startup****BackTalk*** Tarun Mahadevan, Anusha Pandey, Rowan Sperling, Shriya Tiku*Entrepreneurship Award****Koalaty Care*** Neha Madanagopal, Keeerthana Radapuram, London Siever, Lauren Wetzel*Audience Favorite****Concleat*** Ekansh Arora, Abhinav Dandu, Azhan Mohammed, Luis Padilla*Gallery Walk Audience Favorite***Research Track**

1st Place Shruthika Balasubramanian

*“Probing Demographic Factors and Genetic Signatures in Glaucoma Onset and Progression: A Computational Approach”*2nd Place Sanjana Katti*“Effects of Pharmaceuticals and Ocean Warming on Zooxanthellae Expulsion and Asexual Reproduction in Aiptasia: A Model for Tropical Coral”*3rd Place Nitya Bhartiya*“Color Me Surprised: Identifying and Evaluating Logo Color Trends’ Impacts on Business Brand”***STEM Expo Organizing Committee**

Anderson Buchholz, Angelina Durbin, Karina Kejriwal, Rithika Madey,

Roby Matta, Dhruv Naik, Parik Venkataraghavan

Technology Student Association (TSA)*State and Nationals May – July 2024***Technology Student Association (TSA) State****SeaTac, Washington (May 2024)****Animatronics**6th Place Anoushka Gharat, Tabina Hussain, Manas Mishra, Anvie Sana, Manha Zakaria**Audio Podcasting**26th Place Zita Cheung, Sanjana Katti, Siddhesh Nikam27th Place Gabriel Austin, Lila Black**Biotechnology Design**1st Place Raghav Kalyana Sundaram, Sanjana Katti, Saranya Voruganti2nd Place Aaryan Dhingra, Tiana Dumitrescu, Prakshi Shukla, Nikki Taleghani6th Place Likitha Nanduri, Divya Ramu**Board Game Design**2nd Place Zixuan Shang, Sanskriti Sinha, Pallavi Thiruveedu10th Place Arda Keskinpala, Kyle Selman18th Place Alekhya Nidadavolu, Saranya Pammi, Sahaana Sridhar

CAD 2D, Architecture

12th Place Keerthana Radapuram

CAD 3D, Engineering

16th Place Kaya Hurmali

Career Prep

5th Place Hem Akarapu

7th Place Rithika Madey

Children's Stories

11th Place Vivian Li

23rd Place Nitya Bhartiya, Marta Jayantilal, Shriya Thakur

33rd Place Saahiti Mamidipaka, Alekhya Nidadavolu

Coding

5th Place Brenda Li, Divya Ramu

11th Place Kaushal Dabbiru, Aarush Kapoor

14th Place Farzad Hasan, Colin Xie

21st Place Vasudha Narayanan, Aarushi Verma

Concept Art

8th Place Gabriel Austin

Data Science and Analytics

1st Place Shruthika Balasubramanian

Debating Technology Issues

10th Place Anson Chen, Alexander Zhang

15th Place Vivaan Pradhan, Sahaana Sridhar

16th Place Akshara Manchikalapati, Sanskriti Sinha

Digital Video Production

5th Place Gabriel Austin, Lila Black, Zita Cheung, Anvie Sana

5th Place Sahana Kaliyaperumal, Akanksha Revuru, Shreya Santhosh

8th Place Cleofus Balaranjith, Soham Kulkarni, Brian Lee, Demian Orlov, Ryan Park

Drones Challenge

14th Place Ekansh Arora, Saahil Menta, Siddhesh Nikam, Krishnaprasath Senthil Kumar

18th Place Achintya Agrawal, Farzad Hasan, Pranav Kannepalli, Aadhyia Patel

Engineering Design

1st Place Iris Dey, Mehuli Nath, Saranya Voruganti

3rd Place Achintya Agrawal, Saahiti Mamidipaka, Saranya Pammi, Abhinav Vallabhaneni

5th Place Raehan Baliga, Akshay Karnik, Rajvir Mishra

8th Place Shauryan Deb, Aaryan Dhingra, Tiana Dumitrescu, Farzad Hasan

11th Place Muhammad Ahmad, Yatin Mokana, Vivaan Pradhan, Sacchin Saravanan, Krishnaprasath Senthil Kumar, Kirit Singh

Essays on Technology

4th Place Saranya Pammi

Extemporaneous Speech

6th Place Tiana Dumitrescu

11th Place Prarthana Mishra

Fashion Design

20th Place Nitya Bhartiya, Marta Jayantilal

29th Place Sahana Kaliyaperumal, Shreya Santhosh

30th Place Rithika Madey, Dhweya Modi

Forensic Science

1 st Place	Sanjana Katti, Grace Zhao
4 th Place	Vasudha Narayanan, Anjana Punniamoorthy
6 th Place	Megan Halim, Eva Jain

Future Technology Teacher

7 th Place	Akshaya Gowrabathini
-----------------------	----------------------

Geospatial Technology

1 st Place	Aaryan Dhingra, Tiana Dumitrescu, Prakshi Shukla
2 nd Place	Eva Jain, Roxana Stanescu
3 rd Place	Zixuan Shang, Sanskriti Sinha, Pallavi Thiruveedu

Infographic

4 th Place	Sahaana Sridhar
-----------------------	-----------------

Manufacturing Technology

3 rd Place	Layla Kaim, Arlene Shergill
-----------------------	-----------------------------

On Demand Video

1 st Place	Brenda Li, Keerthana Radapuram
13 th Place	Zita Cheung, Arnav Mittal, Pallavi Thiruveedu

Prepared Presentation

Megan Halim – advanced to semi-finals list
Shriya Thakur –advanced to semi-finals list

Promotional Design

4 th Place	Aadhya Patel
9 th Place	Anindita Purkayastha

Software Development

9 th Place	Kaushal Dabbiru, Aarush Kapoor, Ishaan Kothari, Arnav Mittal, Ansh Mohindru
22 nd Place	Hem Akarapu, Leo Pan

STEM Mass Media

2 nd Place	Karma Ibrahim, Shailja Tyagi
-----------------------	------------------------------

Tech Problem Solving

10 th Place	Aashi Juneja, Roxana Stanescu
12 th Place	Vasudha Narayanan, Anjana Punniamoorthy

Transportation Modeling

3 rd Place	Rhea Singh
6 th Place	Anindita Purkayastha

VEX Robotics

2 nd Place	Akshath Dommaraju, Mayank Pandey, Neil Parikh, Sathkrith Segu
3 rd Place	Shauryan Deb, Aaryan Dhingra, Farzad Hasan

Video Game Design

10 th Place	Bhanu Atmakuri, Nitya Bhartiya, Alekhya Nidadavolu, Joel Olson, Swayam Pharate – 10 th Place
11 th Place	Shaurya Gupta, Karthik Rajagopal

Virtual Reality Visualization

2 nd Place	Bhanu Atmakuri, Vivaan Pradhan, Sacchin Saravanan, Krishnaprasath Senthil Kumar
3 rd Place	Raghav Kalyana Sundaram, Pranav Kannepalli, Siddhesh Nikam, Aryaman Rtunjay
5 th Place	Eva Jain, Aashi Juneja, Yusuf Nemutlu, Roxana Stanescu
6 th Place	Kaushal Dabbiru, Inesh Dey, Arnav Mittal, Ansh Mohindru

Webmaster

2nd Place Achintya Agrawal, Shruthika Balasubramanian, Raghav Kalyana Sundaram, Pranav Kannepalli, Aryaman Rtunjay

**Technology Student Association (TSA) Nationals
Orlando, Florida (June 2024)**

Engineering Design

4th Place Iris Dey, Mehuli Nath, Saranya Voruganti

Forensics

7th Place Sanjana Katti, Grace Zhao

Geospatial

6th Place Tiana Dumitrescu, Prakshi Shukla

Software Development

3rd Place Bhanu Atmakuri, Pranav Kannepalli, Vivaan Pradhan, Sacchin Saravanan, Krishnaprasath Senthil Kumar, Abhinav Vallabhaneni

Virtual Reality

10th Place Bhanu Atmakuri, Vivaan Pradhan, Raghav Kalyanasundaram, Aryaman Rtunjay, Sacchin Saravanan, Krishnaprasath Senthil Kumar

Webmaster

9th Place Achintya Agrawal, Shruthika Balasubramanian, Raghav Kalyansundaram, Pranav Kannepalli, Aryaman Rtunjay, Abhinav Vallabhaneni

RECOGNITION:

Defne Dingiloglu, for her research paper, “*The Untapped Potential of the Endophytic Actinomyces, Streptomyces Scabiei, in Producing Antibiotic and Antitumor Secondary Metabolites.*”. This research paper has been accepted and published in *The Journal of High School Science*.

Note: The Journal of High School Science is a peer-reviewed high school science journal, providing a platform for motivated young researchers to engage with the scientific community and gain experience in academic publishing. <https://jhss.scholasticahq.com>