INTERNATIONAL YOUNG RESEARCHERS' CONFERENCE

TOKYO 東京

CONFERENCE PROCEEDNGS 2019

2nd International Young Researchers' Conference
Nov 29th- 30th, 2019
Tokyo, Japan

Proceedings of the 2nd International Young Researchers' Conference (IYRC)
November 29-30, 2019, Tokyo, Japan

Edited by

Aneesh Chand

Published 2019

info@the-iyrc.org

Contents

Organizing Committee

Chairs

Aneesh Chand *Ph.D Robotics, University of Tsukuba, Japan*
Neil Nguyen *M.A Economics, University of Tokyo, Japan*
Varun Sharma *B.Eng Computer Science & Finance, University of Tsukuba, Japan*
Elizabeth Feldeverd *M.S Molecular Biosciences & Bioeng., Uni. of Hawai'i at Mānoa, US*

Review Committee

Abbi Hamed *Ph.D Mechatronics, Imperial College London, UK*
Carolina Alexiou *M.S Mathematics, University of Tokyo, Japan*
Emma Madden *B.A Political Science, University of Chicago, US*
Hayden Mueller *B.A German Studies and Theater & Dance, Trinity College, US*
Marco Mollinetti *M.S Computer Science, University of Tsukuba, Japan*
Michael Amano *B.A Neuroscience and Asian studies, Bowdoin College, US*
Phong Nguyen *M.S Information Technology, Carnegie Mellon University, US*
Rachel Chiu *B.A Gender Studies and Psychology, Harvard University, US*
Subodhana Wijeyeratne *M.A History, Harvard University, US*
Victor Torres *M.S Computer Science, Tokyo Institute of Technology, Japan*

US Advisor

Paul Lewis *B.Sc Neurobiology, Harvard College Vision, US*

Opening

Welcome Speech

Dear Participants,

In 2017, a group of concerned tutors, mentors and academics discussed the need for a research conference for high school students. Out of that discussion, the idea for the International Young Researchers' Conference (IYRC) was formed.

The inaugural 1st International Young Researchers' Conference was therefore held in November of 2018. It was a great success. 24 submissions from 7 different schools in 3 countries were received. Presentation topics ranged from the cancer-fighting properties of natto to machine learning, biofuels and smart soccer balls. Keynote speakers from international and Japanese schools, the National Cancer Research Institute and the Harvard Global Health Leadership Conference also joined forces to encourage more high school students to conduct proper scientific research and to publish and present their results to their peers. The benefit of organizing the conference became apparent within the first hour when it was seen that students from different schools were researching similar topics; the conference created an opportunity for them to share their results with each other.

We are pleased to re-convene the 2nd International Young Researchers' Conference. It aims to bring together school students around the world working on research projects in order to let them share results with their peers.

Submission Statistics

Statistics for this year's conference are as follows: A total of 26 papers were received from 15 schools in 5 countries initially. From these, 20 papers were published. Compared to 2018, IYRC is more diverse, as shown in Fig. 1.

All papers were reviewed by at least two reviewers from the Technical Committee. The reviewers are all highly qualified academics and researchers coming from top universities

across the world. This year we have two keynote speakers, one each from industry and academia.

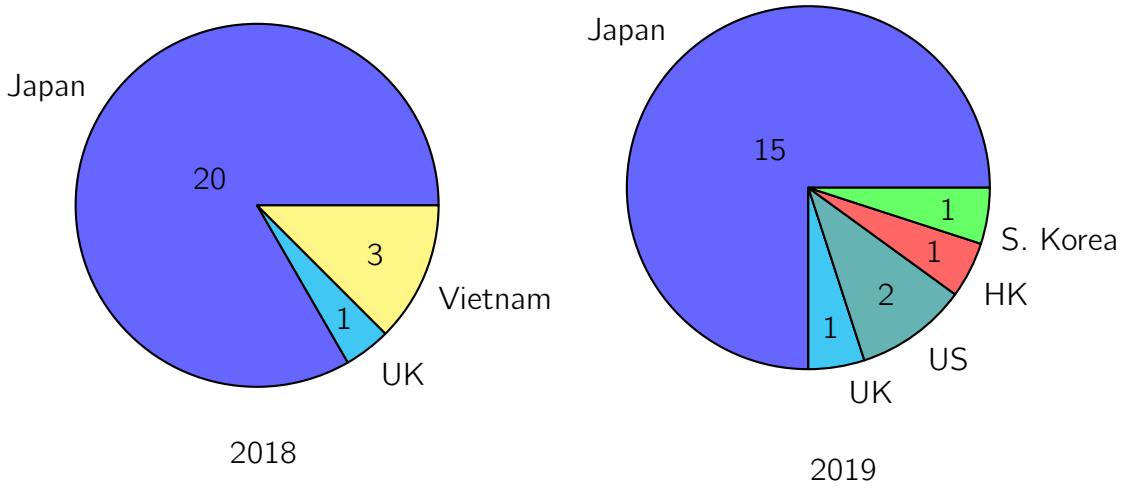


Figure 1: Submission comparison of 2018 and 2019

We acknowledge the management and staff of Tokyo Academics and Tokyo Techies for their sponsorship and support for this event. They have kindly provided us with the use of their premises to host this event. We are grateful to our keynote speakers: Mr. John Mackenzie and Mr. Chris Hartz for kindly accepting our invitation to speak to you and taking out their time to attend this conference.

To participants coming from overseas, Tokyo extends a warm welcome to you.

We hope that that the young minds present here today benefit from this event.

IYRC Chairs

Program

Conference program at a glance

Time	Friday, Nov 29	Saturday, Nov 30
09:00–09:45	Registration	
09:45–10:00	Opening Speech	Registration
10:00–10:45	Keynote speech 1	Keynote speech 2
10:45–11:00	Break	
11:00–13:00	Student Presentations (see pp. 7 for details)	Student Presentations (see pp. 8 for details)
13:00–14:30	Lunch	
14:30–16:30	Student Presentations (see pp. 7 for details)	Student Presentations (see pp. 8 for details)
16:30–17:00	Free Time	Awards and Closing Ceremony
17:00–18:00	Free Time	
18:00–20:00	Dinner	

Each presenter will have 15 min. of presentation time + 3 min. for Q&A.

Awards

The following awards will be presented:

- **Best Paper (STEM)**
- **Best Presentation (STEM)**
- **Honorable Mention (STEM)**
- **Best Paper (Humanities / Social Sciences)**
- **Best Presentation (Humanities / Social Sciences)**
- **Honorable Mention (Humanities / Social Sciences)**

Keynote Speakers

Chris Hartz

Chris Hartz is the General Manager and Representative Director of Mallinckrodt Pharma K.K. where he manages numerous clinical trials and development projects to gain clinical approval and access to ethical drugs and devices for Japanese patients. He has 25 years of experience in the MedTech and Biotechnology fields, with the majority spent in, or working with, Japan. Mr. Hartz has a Bachelor of Arts degree from the University of California, Riverside and a Masters Degree in International Management from Baylor University.

John Mackenzie

John Mackenzie is the Head of the Science department and Research Coordinator within the International Baccalaureate department at Ritsumeikan Uji High School in Kyoto. He is an educator with over 10 years of career experience and research in the field of teaching and learning in the UK and Japan. Mr. Mackenzie graduated with a BSc Honours degree in Aerospace and Aviation Technology from the University of Leeds. He also completed his Post-Graduate Certificate of Education and obtained a Master's Degree in Education.

List of Presentations

Friday, Nov 29th

Time	pp	Name School, Country	Title
11:00–11:18	13	**Kodai Kobayashi** Ritsumeikan Uji Junior and Senior High School, Japan	The Effect of Temperature on the Cross-linked Immobilized Lactase Activity
11:18–11:36	19	**Skyler Pang** St. Mary's International School, Japan	Exercise Recommendations for Patients with Thalassemia Major
11:36–11:54	25	**Yuki Agarwala** K International School, Japan	The Relationship Between the Consequences of Climate Change and Methanogenic Archaea
11:54–12:06	**Break**		
12:06–12:24	31	**David Bass** The American School in Japan, Japan	Comparative Analysis of IMUs and Motion Detecting Technology for Measuring Paitents' Lumbar and Cervical Spines' Range of Motion
12:24–12:42	37	**Edison Suzuki** The American School in Japan, Japan	Study on a Method of Resolving Common Misjudgment by Color Sensors of Robots used in Education
12:42–13:00	41	**Noemie Voss** Sevenoak's School, UK	End-to-end Classification of Ballroom Dancing Music Using Machine Learning
13:00–14:30	**Lunch**		
14:30–14:48	46	**Rainata Vania Hemawani** Ritsumeikan Uji Junior and Senior High School, Japan	Classical Cryptography Methods: Security Improvements by the Hill Cipher
14:48–15:06	52	**Shreya Subramaniam** The American School in Japan, Japan	Text Classification Based on Machine Learning: Quantifying the Sentiment of Online Reviews

15:06–15:24	56	**Airi Ueda** Ritsumeikan Uji Junior and Senior High School, Japan	How can Expedia Group Inc. effectively use HomeAway brand to maintain its position as the market leader with Airbnb Inc. rapidly expanding their Online Travel Agency business?
15:24–15:36	**Break**		
15:36–15:54	62	**Kazuma Takahashi** Ritsumeikan Uji High School Japan	Rebranding of General Motors
15:54–16:12	68	**Jaewoo Lee** Ritsumeikan Uji Junior and Senior High School, Japan	Understanding History Flow of Mathematics from History of Basel Problem
16:30–18:00	**Free Time**		

Saturday, Nov 30th

Time	**pp**	**Name** School, Country	**Title**
11:00–11:18	72	**Kaito Nakamichi** Ritsumeikan Uji Junior and Senior High School, Japan	Deriving a Function Which Expresses the Velocity of an Object Rolling Down a Complex Surface
11:18–11:36	78	**Mayuko Yamamoto** Ritsumeikan Uji Senior High School, Japan	An Investigation Into the Sodium-metal Battery As an Alternative For Lithium-ion Batteries In the Face of Rising Global Demands For Lithium Shortage
11:36–11:54	84	**Sosuke Kojima** Ritsumeikan Uji Junior and Senior High School, Japan	An Analysis On The Effect Of Changing Propeller Diameter And Pitch On The Propulsive Efficiency
11:54–12:06	**Break**		
12:06–12:24	90	**Anna Toneva** Prospect High School, USA	Comparison Between Healthcare Systems and the United States

12:24–12:42	94	**Emmalyn Kartchner** Yokota High School, Fussa, Japan	The Short-Term Effects of Sugar on Self-Worth
12:42–13:00	100	**Justine Kum** Dulwich College Seoul, S. Korea	Social Media and its effects on the Foreign Language Curriculum
13:00–14:30	**Lunch**		
14:30–14:48	107	**Pranav Kanmadikar** DuPont Manual High School, USA	Equalizing Energy In The Food-Energy-Water Security Nexus: An Integrated Global Health Approach Using Reconstructed Assessment And Bioenergy Experimentation
14:48–15:06	113	**Winston Jim Lam** West Island School, HK	Reduction in air pollution to build and maintain a sustainable global society
15:06–15:24	119	**Kei Yamashita** American School in Japan, Japan	Authorship Attribution of 50 Victorian Era Novelists with Convolutional Neural Networks
15:24–16:30	**Break**		
16:30–17:00	**Awards and Closing Ceremony**		

List of Participants

Airi Ueda	Ritsumeikan Uji Junior and Senior High School, Japan
Anna Toneva	Prospect High School, USA
David Bass	The American School in Japan, Japan
Edison Suzuki	The American School in Japan, Japan
Emmalyn Kartchner	Yokota High School, Fussa, Japan
Jaewoo Lee	Ritsumeikan Uji Junior and Senior High School, Japan
Justine Kum	Dulwich College Seoul, S. Korea
Kaito Nakamichi	Ritsumeikan Uji Junior and Senior High School, Japan
Kazuma Takahashi	Ritsumeikan Uji High School, Japan
Kei Yamashita	American School in Japan, Japan
Kodai Kobayashi	Ritsumeikan Uji Junior and Senior High School, Japan
Mayuko Yamamoto	Ritsumeikan Uji Senior High School, Japan
Noemie Voss	Sevenoak's School, UK
Pranav Kanmadikar	DuPont Manual High School, USA
Rainata Vania Hemawani	Ritsumeikan Uji Junior and Senior High School, Japan
Shreya Subramaniam	The American School in Japan, Japan
Skyler Pang	St. Mary's International School, Japan
Sosuke Kojima	Ritsumeikan Uji Junior and Senior High School, Japan
Winston Jim Lam	West Island School, HK
Yuki Agarwala	K International School, Japan

Extra Information

Your **Conference Pack** includes your copy of this conference proceeding and your Certificate of Participation. Please make sure you have these.

Lunch will be served here in the conference venue at 13:00

Wi-Fi will be available during the conference.

The **Conference Dinner** will be held at the "On-Yasai" restaurant. You may either go to the restaurant venue directly (Address: Roppongi 7-13-6, Kondou Building 2F/3F) using the map below or meet outside this building at 17:45.

Contributed Papers

Effect of Temperature on Cross-linked Immobilized Lactase Activity

Kodai Kobayashi
Ritsumeikan Uji Senior High School, Kyoto, Japan
Email: Kobayashi.kodai02@gmail.com

***Abstract*- This paper explores an effect of cross-liked immobilization on the thermostability of lactase enzyme. It is known that the critical denaturization temperature for free-lactase is around 40℃ at pH 7. The paper reports a significant increase in thermostability of lactase (about a 10℃ temperature increase at pH 7) due to cross-linked immobilization based on Tukey HSD Post Hoc ($p < 0.05$). Although further experimentation is needed, it can be suggested that the immobilization of lactase could be used as a viable method to increase the industrial efficiency for production of lactose-free milk, due to the increased reaction rate and reduced risk of microbial contaminations at the elevated operation temperature.**

Key Words – lactase intolerance, immobilization, cross-linking, thermostability, industrial efficiency

INTRODUCTION

Lactose hydrolysis by lactase is an important biotechnical process especially for people with lactose-intolerance. [1] Although I personally love consuming dairy products, I have seen family members and friends who cannot drink as much milk as they want due to lactose intolerance. In fact, lactose intolerance is very common in East Asia including my country, Japan, affecting more than 90 percent of adults.[2] Therefore, I became interested in investigating about the manufacturing procedure of lactose-free milk. The immobilization of lactose is a crucial biotechnology tool as it enables the reusing of enzyme and increases its resistance against chemical and thermal inactivation. [3] I focused particularly on the increased thermal stability of immobilized lactase as it brings numerous commercial benefits to dairy processing industry. Producing lactase-free milk at higher temperature allows for the increased reaction rates, reduced operation time and reduced risk of microbial contaminations. [4] This could potentially optimize the manufacture process of lactose-free milk, thereby helping people with lactase-intolerance to enjoy dairy products.

BACKGROUND RESEARCH

Lactase, also known as beta-galactosidase, is enzyme responsible for breaking lactose into glucose and galactose through a hydrolysis reaction.[5] The reaction is a temperature-dependent reaction because the enzyme activity increases as the temperature increases until it reaches the critical denaturization temperature. The increase is attributed to i) more energetic collisions between particles to overcome the activation energy required for successful collisions ii) more frequent collisions between substrates and active cites due to faster molecular motions. [6] The critical denaturization temperature for free-lactase is known to be around 40℃ at pH 7[7]. After the denaturation temperature is reached, lactase activity decreases as the active sites of lactase deforms[8]. This is caused by the confirmation changes of lactase due to excessive heat[8].

The confirmation stability of enzyme depends upon stabilizing forces arising from the interactions within a protein structure.[9] A cross-linking method of immobilization is an effective technique which results in the thermally and chemically stable immobilization. This is because it increases conformational stability and reduces conformational flexibility of enzyme with additional covalent intermolecular cross-links which are chemically and thermally stable than that of the original stabilizing forces of enzyme.[10]

Glutaraldehyde is widely used as an effective cross-linker reagent which reacts with amine groups of the enzyme.[11]

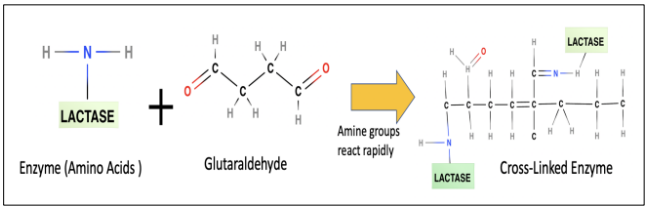


Figure 1: A cross-linking reaction through glutaraldehyde[12]

[1] F. F. Freitas, 2011
[2] US Government, 2018
[3] Andrew Allott, 2014
[4] Claire Vieille and Gregory J. Zeikus, 2001
[5] Quinn Z.K Zhou and Xiao Dong Chen, 2001
[6] Alessandra Bosso
[7] Trenton W. Horner, 2010
[8] Roy M. Daniel and Michael J. Danson, 2013
[9] Daniel, R M, 1996
[10] D. N. Brems, 1991
[11] Isabelle Migneault, 2004
[12] MolView, 2015

A potential drawback of the cross-linking method via glutaraldehyde is the risk of causing the conformational changes of the lactase due to its reactivity and acidity, which reduces the lactase activity. [13] Therefore, glutaraldehyde concentration must be kept between 1 to 5% and a pH buffer is required to maintain the neutral pH, preventing the conformational changes of the enzyme.[14]

HYPOTHESIS:

1) The optimum temperature for cross-linked immobilized lactase activity is expected to be higher than 40℃ (the critical denaturization temperature for the free-lactase) because the cross-linked immobilized lactase has higher conformational thermal stability than that of free lactase[10].

2) The rate of denaturation of cross-linked immobilized lactase is expected to be slower than that of the free lactase due to its conformational thermal stability[10]

MATERIALS/EQUIPMENT

- Digital mass balance: ±0.01g
- Glucometer, sensor strips: ±5mg/ml
- pH strips
- Water bath:
- Thermometer: ±1℃
- Graduated cylinders: 300 mL
- glass pipettes: 100mL
- Tea strainer: none
- Stirring rods, magnetic stirrers
- Bunsen burner, matches tripod, gauze mat
- Separable funnels with a stopcock
- Refrigerator
- Calcium chloride ($CaCl_2$): 6.65g
- Sodium alginate ($C_6H_9NaO_7$): 2.2g
- Glutaraldehyde ($C_5H_8O_2$): 50 mL
- 0.5M of phosphate buffer (pH 6.88): 2000mL
- Lactose supplements: 1g
- Whole milk: 290 mL
- Distilled water: 200 mL

Immobilization: Physical Absorption Method:[15]

1.1. Prepare 100mL of a 0.5M $CaCl_2$ solution
1.2. 2.2g sodium alginate is dissolved in 100mL of 0.5M of phosphate buffer (pH 6.86 due to heat)
1.3. Heat the solution up in order to dissolve sodium alginate
1.4. Add and mix 1g of lactase into alginate hydrogel (make sure that sodium alginate has cooled down as adding enzyme in a high temperature solution may cause enzyme denaturization)
1.5. Mold alginate-lactase mixture with a spike and a stirring rod
1.6. drop a sodium alginate solution into 0.5M of $CaCl_2$ solution, forming tear-drop shape alginate beads with lactase
1.7. Sieve off the alginate beads from the solution with a tea strainer.
1.8. Stored beads under in 200mL of 0.02M $CaCl_2$ solution overnight at 4℃ for hardening

2. Immobilization of Enzyme: Cross-linking Method:[16, 17]

2.1. Add 10 mL of glutaraldehyde and 10 grams of beads in 323 mL of phosphate buffer
2.2. Expose the mixture to magnetic stirring for 3 hours for cross-linking
2.3. Rinse the beads with distilled water to remove excess glutaraldehyde

3: Using a Funnel Bioreactor Set-up:[18,19]

3.1. Place 3 g of cross-linked immobilized lactase beads in a separating funnel
3.2. Let the separable funnel sit in a water bath at the desired temperature for 3 minutes
3.3. Pour 6 mL of milk into the funnel
3.4. Time for 2 minutes for lactase to carry out its reaction in a water bath
3.5. Open a stopcock, allowing milk to flaw down to a beaker on the bottom
3.6. The beads are replaced every 3 times of use (2 times for the last 4 and 5 trials)

4: Measuring Lactase Activity:[20]

4.1. Place a drop of the resultant milk onto a tip of a glucometer sensor strip
4.2. A total of 5 trials are carried out for each temperature

Experimental Procedure Snapshots:

Figure 2: *Alginate Beads*

Sodium alginate was molded into a tear-drop shape using a spike and a stirring rod. Alginate beads before the cross-linking phase. All beads are opaque and has tear-drop shape.

[13] Sumitra Datta, 2013
[14] He Chen, 2013
[15] Sachin Talekar and Sandeep Chavare, 2012
[16] Bharat Bhushan, 2015
[17] BioNetwork, 2013
[18] The El Paso County Community College
[19] Nam Sun Wang
[20] Sanwa Chemical research institute Japan, 2018

Figure 3: Water bath and the bioreactor

RISK ASSESSMENT

The risks associated with calcium chloride and glutaraldehyde listed above were minimized by working with lower concentrations of these solutions. However, safety goggle and gloves were stilled used as a protection. A Bunsen burner is placed away from combustible materials, and it is adjusted to produce a medium blue flame. The gas is shut off when its use it complete. The chemicals are environmental hazards if disposed with common water drain. Each chemical was disposed separately as either general waste (calcium chloride and buffers) and organic waste (sodium alginate). Products can also be disposed as organic waste.

DATA COLLECTION

Although it is be idealistic to account for an initial glucose concentrations and pH of milk prior to every trial, this was not feasible. This was because I could not afford to buy sensor strips to account for every trial as they were expensive and laboratory time was limited. The initial concentration and pH were measured only 5 times before the reaction as shown on Table 1.

	Trial 1	Trial 2	Trial 3	Trial 4	Trial5
The Intial Glucose Concentrations (mg/dL) (±5 mg/dL)	>10mg/dL	>10mg/dL	>10mg/dL	>10mg/dL	>10mg/dL
pH Level	Netural	Netural	Netural	Netural	Netural

Table 1: The Initial Glucose Concentrations and the pH of Milk

The glucose concentration was measured by glucometer, whereas pH was measured by pH strips. Glucose concentrations lower than 10 mg/dl are indicated as low since the glucometer cannot measure a concentration lower than 10mg/dl. Overall, the result suggests that the initial glucose concentrations of milk were all consistently lower than 10mg/dl and around pH 7 respectively, suggesting that the initial condition likely to have stayed the same through the experiment.

In addition to this, the experiment was carried out with alginate beads not containing any lactase enzyme (instead of the cross-linked lactase alginate beads) as a control. This was carried out only once for each independent variable due to the same reason described earlier.

Temperature/℃ (±1℃)	Glucose Concentrations (mg/dL) (±5 mg/dL)
4	>10mg/dL
15	>10mg/dL
40	>10mg/dL
45	>10mg/dL
50	>10mg/dL
60	>10mg/dL
80	>10mg/dL

Table 2: Data collected from the controlled experiment (with alginate beads not containing lactase enzyme)

The result indicates that the glucose concentration of milk did not change from its initial glucose concentration (low), suggesting that temperature nor alginate beads (no lactase) has an effect on the glucose concentration of milk. Thus, this suggests that the cross-linked immobilized lactase in the beads are likely to be a factor which affected the glucose concentration of milk.

Table 3 shows the effect of the temperature on the lactase hydrolysis activity, measured by the glucose concentrations of the resultant milk.

	Glucose Concentrations (mg/dL) (±5 mg/dL)		
Temperature/ ℃ (±1℃)	Mean	Maximum	Minimum
4	63	72	55
15	114	133	100
40	171	178	161
45	200	208	193
50	295	300	291
60	251	269	238
80	102	108	98

Table 3: The Mean and Range of the Data

DATA ANALYSIS

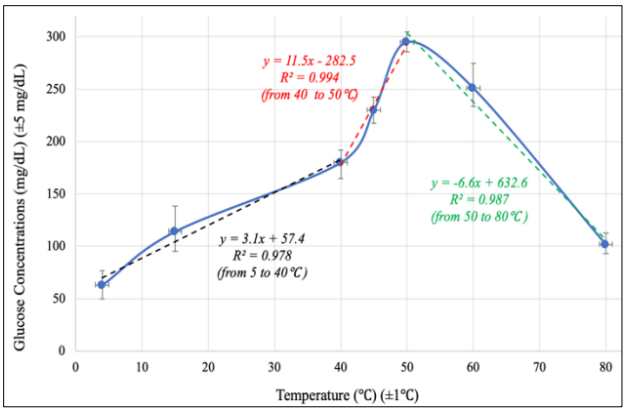


Figure 4: The effect of the temperature on the hydrolysis reaction by cross-linked immobilized lactase measured by the glucose concentration of milk (with the mean of each independent variable)

Overall, the cross-linked lactase activity shows a positive trend up to its critical denaturization temperature as the glucose concentration increases from 4℃ up to 50℃ indicated by the positive slopes. The rates of increase, however, are inconsistent as a gradual increase is seen between 4 to 40℃ indicated by a slope of 3.1x, whereas a sharp increase is observed between 40 to 50℃ indicated by a slope of 11.5x. The concentration of glucose in the milk reached its optimum of around 300mg/dl at 50℃. There is a steady decrease from the optimum point onward indicated a negative linear slope of $-6.6x$. A high R_2 values for each line: 0.978, 0.994 and 0.987 suggest that a linear regression line is a suitable type of line for all intervals.

The green trend line specifically focuses on the denaturing rate of the cross-linked lactase. From the trend line, it is deduced by taking a derivative of the slope that a rate of denaturation of the cross-linked lactase was -6.6. This implies that the glucose concentration of milk (the lactase activity) decreases by approximately 6.6mg/dl for every increase of 1℃.

Although the non-overlapping consecutive range bars are seen, statistical analysis (ANOVA and Post Hoc tests) are carried out in order to confirm that the difference is statistically significantly different. Although it must be noted that 7 samples are not sufficient to draw a statistically significant analysis and conclusion, statistical analysis shown below follows a valid procedure and should be considered as a valid analysis if sufficient samples are to be collected. However, more samples must be collected for improvement.

STATISTICAL ANALYSIS

In order for parametric statistical tests such as *one-way ANOVA* and *Tukey HSD Post Hoc* to be valid, three assumptions listed below are needed to be met[21]. These assumptions include *"independence of observations assumption*" (observations are independent of one another), "*normality assumption*" (data are normally distributed in each of the treatment) and "*homogeneity of variances assumption*" (the variances are equal across treatment group). My data satisfies the first condition because the observations are independent of one another as each trial is conducted independently of other trials.

Shapiro-Wilk test is used to confirm normality assumption, while Bartlett's test is used to confirm the homogeneity of variance assumption for the data sample.

Normality Assumption (Shapiro-Wilk Test)[22]:

Shapiro-Wilk test is conducted to confirm that my data satisfies the normality assumption.

H_0: The data set is normally distributed.

H_a: The data set is not normally distributed.

Group (℃)	skewness	Excess kurtosis	W Statistic	p-value	Critical Value	Accept/reject the alternative hypothesis
4	0.191	0.06	0.998	1.000	p <0.05	Reject
15	0.550	-1.46	0.942	0.844	p <0.05	Reject
40	-0.802	-0.80	0.921	0.678	p <0.05	Reject
45	0.076	-2.55	0.920	0.669	p <0.05	Reject
50	0.494	-3.16	0.794	0.098	p <0.05	Reject
60	0.554	-0.74	0.923	0.694	p <0.05	Reject
80	0.863	1.09	0.956	0.939	p <0.05	Reject

Table 4: Shapiro-Wilk Test

Each group is tested separately as each sample is assumed to be taken from an independent population. The result above rejects the alternative hypothesis and accepts the null hypothesis. Therefore, my data satisfies the normality assumption.

Bartlett's test is used to confirm that my data satisfies the homogeneity of variance assumption.

H_0: The data has an equal variance across the group.

H_a: The data does not have equal variance across the group.

Group (℃)	4	15	40	45	50	60	80
Sample Variance	40.7	186.7	47.5	43.3	19.7	165.3	14.2
x^2	10.506						
p-value	0.105						
Critical Value	p-value < 0.05						

Table 5: The Bartlett's Test: Equality of variances

Since p-value of 0.105 > 0.05, it rejects the alternative hypothesis and it accepts the null hypothesis of equal variance. Thus, my data satisfies the homogeneity of variance assumption. Since the data passes both of the tests, *ANOVA* and *Tukey HSD Post Hoc* can be used to account for significance of the mean differences.

It is confirmed the data collected satisfies all the three assumptions required to conduct one-way ANOVA and the Tukey HSD. One-way ANOVA test is conducted using excel.

H_0: There is no significant difference in the data set.
H_a: At least 1 data set is significantly different from other data sets.

Source of Variation	*Sum of Square*	*Degrees of Feedom*	*Mean Sqaure*	*F-Statistic:*	*P-value*	*F-Critical:*	*Alpha*
Between Groups	211033	6	35172	476	7.82E-27	2.45	0.05
Within Groups	2070	28	74				
Total	213103	34					

Table 6: One-Way ANOVA Test Calculation

Since p-value of $7.82 \times 10^{-27} < 0.05$, it rejects the null hypothesis and accepts the alternative hypothesis that there are at least 1 statistically significant difference in the data set. However, ANOVA test is used to test general rather than specific differences. Therefore, the Tukey HSD was used to compare two specific means and determine if a difference

[21] Robert M. Pirsig

[22] Admin

between the two means is statistically significant. Since the scope of my investigation is to find whether or not there is a significance increase in the optimum temperature of the cross-linked enzyme from that of the free lactase (40℃), the comparison specifically focuses around the optimum points.

H_0: There is no significant difference between the two groups.
H_a: There is a significant difference between the two groups.

Group 1 Temperature/ ℃ (±1℃)	Group 2 Temperature/ ℃ (±1℃)	$\lvert(\bar{x}_{G1} - \bar{x}_{G2})\rvert$	N_{G1}	N_{G2}	MS_{within}	$\sqrt{\frac{MS_{Within}}{n}}$	Q-Score for p=0.05	HSD Crtical Value
40	45	29	5	5	73.90	3.84	4.45	17.11
40	50	124	5	5				
45	50	95	5	5				

Table 7: Tukey HSD Post Hoc Test

For a difference between two data to be statistically significant, a difference between two means must be greater than the HSD critical value. Since mean differences for each group (29, 124 and 95) is greater than the HSD critical value of 17.11, it rejects the null hypothesis and accepts the alternative hypothesis that there is a statistically significant increase in the glucose concentration between from 40 to 45℃ group, 40 to 50℃ group and 45 to 50℃ group.

Although it must be noted that

EVALUATION

The major limitation was that the experiment was not conducted with free lactase, as opposed to cross-linked immobilized due to the time constrains, which would have strengthened the analysis and conclusion. Although the increase in thermal stability was deduced from the literature value, the comparison between the two made may be insufficient as different lactase could have different initial thermal stability[23]. Moreover, even though I was able to find a denaturing rate of the cross-linked lactase, I could not compare with that of free-lactase as other scientific literatures were based on different units. All in all, comparing my results for cross-linked immobilized lactase with free lactase from my own data would have led to a more valid and clear analysis and conclusion. Furthermore, another valid problem related to this was that I could not confirm whether or not the lactase has successfully been cross-linked as it is impossible to visually assess the successful cross-linking of lactase. Since the successful cross-linking can be deduced from the increased thermal stability of lactase, comparing my own data for both free and immobilized lactase would have been a clear and valid indicator for this.

Moreover, it was difficult to control the temperature using the water bath, which led to uncertainty in the independent variable. I had a trouble holding the separable funnel under water for a long time using a tweezer as it was not stable and was dangerous especially when working with water at high temperatures. Attaching a weight to separable funnel to keep it under water would solve the problems and can even quicken the procedure as it removes a necessity to hold the funnel manually. Furthermore, it was difficult to keep the temperature of water bath the same especially at high temperature such as 50, 60 and 80 ℃ due to low room temperature. This can have been improved using air-conditioning to keep the temperature of room reasonably high and also by covering up the water bath with a glass cover. This would reduce a temperature difference between water bath and its surroundings, thereby reducing the rate of heat transfer between the two environments in order to keep the temperature more consistent.

The results of my experiment are limited as it examines a wide range of temperature. As an extension, it is essential to look at close ranges from between 45 to 60℃ and to identify where exactly the optimum point lies. This would help me pinpoint the optimum temperature of the cross-linked immobilized lactase which can be used to maximize the production of lactose-free milk. However, following my methodology would not work without major refinements. When working with smaller range of temperature, increasing measurement precision and reducing random errors are necessary to attain values precisely closer to the exact glucose concentration to identify the optimum point. Since the glucometer and the water bath has the uncertainty of ±5 mg/dl and ±1℃ respectively, they could be significant for smaller ranges, therefore, invalidate my data**.** Therefore, using more accurate method for controlling temperature and more accurate glucometer are necessary as relative uncertainty will be larger when dealing with smaller ranges.

DISCUSSION/ FINDINGS

With regarding to my second hypothesis, no conclusion can be drawn. Although I found a denaturing rate of the cross-linked immobilized lactase, I could not find a denaturing rate of free-lactase from the literature to compared against as other studies used different units. However, my first hypothesis was supported because the observed optimum temperature of the cross-linked lactase was 50 ℃ at pH 7, which is 10 ℃ higher than that of the commercial free-lactase (40 ℃).[6] I can be reasonably confident of my findings as the increase from the optimum temperature of the free-lactase was shown to be statistically significant by non-overlapping range bars and the Tukey post hoc test**.**

This is further supported by accepted scientific studies which states that the cross-linked enzymes are thermal stability due the additional covalent intermolecular cross-links that are chemically and thermally stable[10]. Similar results were found by Journal of Molecular Catalysis B as the optimum temperature of lactase shifted from 35 to 50 ℃ due to glutaraldehyde binding immobilization**.**[24] According to my findings, it can be suggested that the cross-linked lactase has

[23] NutraTech

[24] J. Szczodrak, 2000

the higher thermal stability than free-lactase, although it would have been best to compare the optimum temperature of the cross-linked lactase against free-lactase from own primary data.

CONCLUSION

Lactose hydrolysis by lactase is an important biotechnical process which could potentially make the processing of lactose-free milk more efficient. This would be highly beneficial for those with lactose-intolerance. Based on my experiment, it can be suggested that the cross-linking lactase can be used to optimize the manufacture process of lactose-free milk by increasing its thermal stability, although further experiment and research are required to confirm this.

ACKNOWLEDGMENTS

The authors wish to thank Mr. Schwarz, my biology teacher, and laboratory technicians for their help and guidance in conducting the experiments and proofreading this paper.

REFERENCES

[1] Admin. “Shapiro-Wilk Test Calculator.” *Statology*, Admin, 28 Jan. 2019
[2] Bergwerf, Herman. “MolView.” *MolView*, AGPL , 2015,
[3] Bhushan, Bharat, et al. “Improved Enzyme Catalytic Characteristics upon Glutaraldehyde Cross-Linking of Alginate Entrapped Xylanase Isolated from Aspergillus Flavus MTCC 9390.” *Enzyme Research*, Aug. 2015,
[4] Bosso, Alessandra et al. *Lactose Hydrolysis Potential and Thermal Stability of Commercial β-Galactosidase in UHT and Skimmed Milk*, Food Science and Technology Campinas
[5] Brems, D N, et al. “The Conformational Stability and Flexibility of Insulin with an Additional Intramolecular Cross-Link.” *Current Neurology and Neuroscience Reports*, U.S. National Library of Medicine, 25 Jan. 1991
[6] Chaplin, Martin. “Methods of Immobilisation.” *Enzyme Technology*, London South Bank University, 2004
[7] Chen, He, et al. “The Effect of Glutaraldehyde Cross-Linking on the Enzyme Activity of Immobilized β-Galactosidase on Chitosan Bead.” *Advance Journal of Food Science and Technology*, vol. 10, no. 7, 2013, pp. 932–935.,
[8] Daniel, R M, et al. “The Denaturation and Degradation of Stable Enzymes at High Temperatures.” *National Center for Biotechnology Information*, Biochemical Journal, July 1996
[9] Daniel, Roy M., and Michael J. Danson. “Temperature and the Catalytic Activity of Enzymes: A Fresh Understanding.” *FEBS Letters*, Elsevier, 27 June 2013
[10] Dittami, Simon. “Shapiro-Wilk Normality Test.” *Shapiro Wilk Test (Web Version)*, Research Gate Scientific Network
[11] “Effect of Temperature on Enzyme Activity.” *Lecture Outline*, City University of New York, academic.brooklyn.cuny.edu/biology/bio4fv/page/enz_act.htm.
[12] “Enzymatic Activity of Lactase.” *Educational Links*, Pottsgrove School District,
[13] “Enzymes.” *Enzyme Lab*, The El Paso County Community College,
[14] Freitas, F. F., et al. “OPTIMIZATION OF THE IMMOBILIZATION PROCESS OF β-GALATOSIDADE BY COMBINED ENTRAPMENT-CROSS-LINKING AND THE KINETICS OF LACTOSE HYDROLYSIS.” *Brazilian Journal of Chemical Engineering*, vol. 29, no. 01, Aug. 2011, pp. 12–24.,
[15] Lactose Intolerance - Genetics Home Reference - NIH.” *U.S. National Library of Medicine*, National Institutes of Health - US Government, 20 Nov. 2018
[16] Lane, David M. *Introduction to Analysis of Variance*, Rice University (Lead Developer), University of Houston Clear Lake, and Tufts University
[17] Migneault, Isabelle, et al. “Glutaraldehyde: Behavior in Aqueous Solution, Reaction with Proteins, and Application to Enzyme Crosslinking.” *BioTechniques*, BioTechniques, Nov. 2004,
[18] Pirsig, Robert M., et al. *Checking Assumptions of One-Way ANOVA*, Australian Center for International Agricultural Research
[19] Saleh Zahediasl, Asghar Ghasemi. “Normality Tests for Statistical Analysis: A Guide for Non-Statisticians.” *Normality Tests for Statistical Analysis: A Guide for Non-Statisticians*, Int J Endocrinol Metab, 20 Apr. 2012
[20] Szczodrak, J. “Hydrolysis of Lactose in Whey Permeate by Immobilized β-Galactosidase from Kluyveromyces Fragilis.” *Journal of Molecular Catalysis B: Enzymatic*, Elsevier, 18 Aug. 2000
[21] Talekar, Sachin, and Sandeep Chavare. “Optimization of Immobilization of α-Amylase in Alginate Gel and Its Comparative Biochemical Studies with Free α-Amylase.” *Recent Research in Science and Technology*, Department of Biotechnology Engineering, Kolhapur Institute of Technology’s College of Engineering, Kolhapur, Maharashtra, India, 2012,
[22] Vieille, Claire, and Gregory J. Zeikus. “Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability.” *National Center for Biotechnology Information*, Microbiology and Molecular Biology Reviews, Mar. 2001
[23] Wang, Nam Sun. “CONTINUOUS IMMOBILIZED ENZYME REACTOR.” *EXPERIMENT NO. 13*, Department of Chemical & Biomolecular
[24] Zhou, Quinn Z.K, and Xiao Dong Chen. “Effects of Temperature and PH on the Catalytic Activity of the Immobilized β-Galactosidase from Kluyveromyces Lactis.” *Science Direct* , Biochemical Engineering Journal, Volume 9, Issue 1, Pages 33-40, 29 Oct. 2001

Exercise Recommendations for Patients with Thalassemia Major

Skyler Pang
St. Mary's International School, Tokyo, Japan
Email: skyler70@smis.ac.jp, pangskyler@gmail.com

Abstract **- Thalassemia major (TM) is an inherited blood disorder caused by one or more genetic mutations that change the structure of hemoglobin and red blood cells (RBCs). This leads to multiple complications including anemia, splenomegaly, excess blood-iron levels, and stunted physical development. Although a growing body of research describes the impact of cardiac iron burden, anemia, and pubertal delay on exercise capacity in thalassemia patients, there is a gap in the clinical literature regarding exercise recommendations for these individuals. Accordingly, this paper evaluates and analyzes relevant literature in order to propose an exercise recommendation model for adolescent male TM patients. It generally concludes that patients should adjust the duration, frequency and intensity of exercise based on the severity of their symptoms.**

Key Words - Thalassemia major, exercise capacity/performance, cardiac iron burden, anemia, puberty delay.

INTRODUCTION

Thalassemia overview

Thalassemia refers to a group of heritable, hemoglobin-related blood disorders [1]. Common in individuals of African, Mediterranean and South-Asian descent, thalassemia stems from an imbalance of alpha and beta globin proteins that make up the hemoglobin molecule [2]. This imbalance impairs hemoglobin protein folding and results in misshapen red blood cells (RBCs), which lead to a range of complications including anemia, enlarged spleen, elevated blood-iron levels, and stunted physical development.

Thalassemia consists of two subtypes, alpha and beta, which indicate whether the genetic mutation occurs in the alpha or beta globin protein [1]. The disease is further classified based on symptom severity as major, intermedia, or minor. This depends on the number and severity of mutations in the four genes encoding alpha globin or in the two genes encoding beta globin [2]. One to two faulty alpha genes cause alpha thalassemia minor, whereas three to four faulty alpha genes presents as the more severe intermedia and major forms of the disease [2]. In contrast, one faulty beta globin gene causes beta thalassemia minor, whereas two faulty beta globin genes can produce either moderate (beta thalassemia intermedia) or severe symptoms (beta thalassemia major, also known as Cooley's Anemia) [2]. Individuals with either major variation require consistent blood transfusions to avoid potentially fatal symptoms; in contrast, minor subtypes do not require transfusions.

Pathophysiology

Thalassemia is characterized by improper folding of the hemoglobin molecule. Hemoglobin is a polypeptide primarily found in RBCs that functions to transport oxygen and carbon dioxide between the lungs and external tissues [3]. Consisting of two alpha and two beta polypeptide chains, each hemoglobin molecule depends on the consistent and balanced production of alpha and beta globin proteins [3]. Alpha globin production is controlled by two genes, *HBA1* and its redundant backup *HBA2*, both of which are located on chromosome 16 [3]. Both of these genes have homologous copies, resulting in four genes that each produce approximately 25% of alpha globin. Accordingly, a mutation in one alpha globin gene that disrupts transcription or renders its translational product non-functional will cause a 25% decrease in alpha globin chains available for use in hemoglobin synthesis. In contrast, beta globin production is controlled by one gene, *HBB*, which is located on chromosome 11 [1]. Each homologous copy of *HBB* produces approximately 50% of beta globin. As a result, even one non-functional copy of *HBB* can severely disrupt beta globin availability for hemoglobin synthesis [1]. Any of these mutations will cause a thalassemia phenotype.

Clinical presentation

TM patients present clinically with a wide range of symptoms including anemia, splenomegaly, and pubertal delay. Depending on symptom severity, each of these can substantially impair an individual's ability to exercise.

Anemia and spleen enlargement

The most prevalent symptom of TM is anemia, or a decreased amount of hemoglobin and/or RBCs [3]. In most cases, a mutation in *HBA1*, *HBA2* or *HBB* causes an imbalance in alpha and globin chain availability. As hemoglobin is produced, excess proteins from the non-mutated, functional gene(s) accumulate inside RBCs or their precursor cells in the bone marrow [3]. This accumulation

has two primary consequences. First, excess globin buildup damages RBCs and their precursors, which frequently causes hemolysis [3]. This triggers hematopoietic stem cell differentiation that further taxes the body's hematopoietic system. The second consequence involves a change in the shape of RBCs. As excess globin chains accumulate, they interfere with RBC's ability to regulate water and ion concentrations [3]. This causes RBCs to morph from a biconcave-disk to a misshapen version with decreased surface area, ultimately inhibiting oxygen's ability to diffuse across the cell membrane and bind to functional hemoglobin [3]. Taken together, these consequences lead to anemia and high levels of cardiovascular stress.

While hemolysis can occur in circulating RBCs, it primarily occurs in the spleen, where abnormal RBCs are filtered through an endothelial slit [4]. In healthy individuals, any misshapen RBCs that are unable to move through this slit are marked for destruction by macrophages. In TM individuals, however, deformed RBCs accumulate faster than they undergo hemolysis. Accordingly, RBCs accumulate causing spleen enlargement and swelling [4].

Transfusion Therapy and Cardiac Iron Burden

In order to combat the harmful effects of anemia, TM patients often receive regular blood transfusions. Together with occasional bone marrow transplants, this helps them to maintain healthy levels of hemoglobin and RBCs [1]. One negative complication of this, however, includes elevated levels of iron throughout the body. Over time, oxidative damage from excess iron levels has been shown to damage the tissue lining of the heart and vasculature, impairing the heart's ability to pump blood and placing a substantial burden on the cardiovascular system [5]. This can cause left ventricular systolic and diastolic dysfunction, which leads to inefficiency in pumping blood and increased risk of heart attack [6]. Prescription of iron-chelating agents has been shown to reduce this damage; however, the extent of reduction varies by individual, and several studies describe harmful consequences from even slightly elevated levels of iron [7][8].

Stunted physical growth

In addition to increased cardiovascular burden, iron overloading has been shown to play a major role in pubertal delay. In 95% of males, pubertal development begins between the ages of 9–14 [9]. This occurs via hormonal signaling in the HPG axis. First, the hypothalamus releases gonadotropin-releasing hormone (GnRH), which activates the release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) from the pituitary gland. These hormones stimulate testosterone (T) production in the gonads, which eventually leads to increases in bone density, muscle mass, and body fat [10]. In the context of exercise, these changes exert protective effects and help to prevent injury.

Several studies have described pubertal delay in TM patients, with negative effects on bone mineral density and muscle mass [7][11]. In one study, patients were shown to have a low spine bone mass density z-score of -2.1 ± 1.1 compared to healthy individuals [11]. This finding is supplemented by other studies that indicate decreased height, weight, and lean body mass [11], in addition to impaired bone formation [7]. Mechanistically, this delay has primarily been attributed to oxidative damage from excess iron levels that disrupts production of LH, FSH, and T [8].

THALASSEMIA AND EXERCISE

The extensive physical and mental benefits of exercise—including weight management, improved cognitive function, increased energy levels, and a more positive outlook on life—are well documented [12][13]. While exercise has been shown to improve patient outcomes in cancer patients, individuals with sickle-cell anemia and other clinical populations [14][15], TM patients have detrimental symptoms that interfere with their ability to exercise. Regular blood transfusions, iron chelation therapy, and side effects of pubertal delay place severe stress on their cardiovascular, hematological, and skeletal systems. Given the risk of injury, death or other disease-related complications, many individuals with TM avoid exercise altogether [11]. Clinicians similarly struggle to recommend a suitable exercise program or even assess a patient's ability to exercise [15]. This is further exacerbated by a gap in the clinical literature regarding exercise recommendations for TM patients. This creates a situation in which TM patients are unable to reap the physical and mental benefits of regular exercise. Accordingly, this analysis seeks to equip physicians with the tools to assess exercise capacity in TM patients and provide suitable exercise recommendations. Specifically, it does this by evaluating TM symptoms as they relate to exercise in the context of previously published data.

Drawing on previous studies, I define exercise capacity as the amount of physical exertion a patient can sustain with regards to intensity, duration, and frequency [15][12]. While there are many ways to define these aspects of exercise, clinical recommendations must provide clear, concrete descriptions in order to prioritize patient safety and adherence [15][12]. Thus, this analysis uses subjective intensities rather than technical measurements (i.e. oxygen pulse, VO_2max) with regards to exercise intensity (Table 1). Duration and frequency were kept consistent with similar recommendation models described in the literature [15][16] (Table 1).

Table 1. Definitions of exercise duration, frequency, and intensity that were applied when constructing an exercise recommendation model for thalassemia major patients.

	Duration [16]	Frequency [16]	Intensity [15]
Low	5–10 minutes	1–2 days a week	Start exercising gradually; stop before the first sign of fatigue
Moderate	~ 20 minutes	3–5 days a week	Start exercising gradually, stop at the first sign of fatigue
High	< 30 minutes	6–7 days a week	Start exercising gradually, stop briefly past the first sign of fatigue

While several studies have described detrimental complications of TM on exercise capacity, this analysis focuses on those symptoms most relevant to exercise. It specifically concentrates on a handful of relevant studies to examine the impacts of iron-cardiac burden, anemia and pubertal delay on the cardiovascular, hematopoietic and skeletal systems, respectively.

Cardiac Iron Burden

In their 2013 study, Sohn et al. [5] investigated the relationship between TM and exercise performance, specifically elucidating the correlation between higher cardiac iron burden and decreased exercise capacity. The experiment consisted of a treadmill exercise stress test, during which seventy-one male and female TM patients of ages 12-45 were told to exercise until exhaustion. Measurements were taken throughout to determine VO_2 max, heart rate (HR), and oxygen pulse. VO_2 max measures maximal oxygen use by an individual during intense exercise [17]. It is described by the Fick equation (1) shown below:

$$VO_2 \text{ max} = Q \times (C_aO_2 - C_vO_2)$$

Here, Q represents the cardiac output of the heart or the volume of blood the heart pumps [17]. C_aO_2 represents the oxygen content in one's arteries while C_vO_2 represents the oxygen content in the veins [17]. In other words, oxygen content in the arteries is subtracted by the oxygen content in the veins to get the amount of oxygen the body uses to function [17]. The Q value for VO_2 max varies considerably based on cardiac iron burden given its direct connection with heart contraction efficiency.

By correlating cardiac iron levels with VO_2 max, HR, and oxygen pulse—a measurement obtained by dividing VO_2 max by HR, the authors demonstrated a significant correlation between elevated cardiac iron levels and decreased exercise capacity. This decrease was primarily driven by a lower VO_2 max, for which the strongest predictor was cardiac iron levels ($r^2 = 0.12$) [5].

While a diminished Q value suggests additional cardiac stress that decreases the efficiency of heart contractions, a decrease in arterial or venous oxygen content reflects issues with oxygen-carrying capacity that are more closely related to anemia. Given cardiac iron burden's status as a stronger predictor of VO_2 max than Hb levels ($r^2 = 0.12$ vs. 0.05), it can be deduced that these exercise-related changes are primarily driven by cardiovascular stress. This could include chronic vascular inflammation, vascular stiffness, and increased rates of cardiac cell death [5][6]. These forms of cardiac stress have been shown to impair exercise ability / put people at risk for heart attack/ etc. [6][18]. Thus, elevated cardiac iron burden emerges as an immediately relevant risk factor to consider when providing exercise recommendations.

Anemia

While the aforementioned study by Sohn et al. [5] highlights cardiac iron burden as one primary factor limiting exercise capacity in TM patients, the authors also note the impact of anemia. When separated into two groups based on Hb level (>12g/dL, "low anemic" and <12g/dL "high anemic"), the authors found a correlation between lower Hb levels and decreased VO_2 max. Specifically, low anemic patients demonstrated VO_2 max that were 83.0 +/- 15.1% of their predicted value. Similar decreases were noted for HR and oxygen pulse, with 93.4+/- 5.5% and 86.6 +/- 14.0% of predicted values observed respectively [5].

These data have further relevance in the context of other hemoglobinopathies in which body tissues do not receive enough oxygen as a result of RBCs' impaired ability to bind and transport oxygen [3]. In one study that sought to evaluate the impact of hemoglobinopathies on exercise capacity [19], the authors removed RBCs from healthy individuals in order to induce anemia and maintained a reduced RBC count for two weeks via regular blood extraction. A comparison of exercise measurements at rest, submaximal and maximal efforts revealed significant decreases in VO_2 max for anemic patients compared to healthy controls, and these decreases were directly proportional to the decrease in oxygen carrying capacity [19]. In another study of sickle cell anemia patients, VO_2 max and HR were observed to be 68% and 86% of their respective predicted values [20]. Collectively, these studies demonstrate the detrimental effects of anemia on exercise capacity in TM patients and provide a framework through which recommendations can be considered.

Exercise Recommendations for Patients with Thalassemia Major

Pubertal delay and growth failure

In addition to cardiac iron burden and anemia, pubertal delay presents a major complication of TM. As bones undergo regular remodeling to replace the entire skeleton, they strike a delicate balance between resorption and formation [21]. Iron overload disrupts this process, leading to bone deformities and decreased bone mineral density. This is further exacerbated by oxidative damage from excess iron that disrupts increases in height, weight and muscle mass that are normally engendered by HPG-axis signaling [8].

While many studies with TM patients have focused on cardiac iron burden and anemia, a growing number of researchers are investigating pubertal delay and its consequences. One important study [8] evaluated pubertal development in thirty-three adolescent males with B-thalassemia between ages 14–21. By injecting these patients and healthy control individuals with an analogue of GnRH, the authors sought to determine the impact of B-thalassemia on HPG axis function. As expected, a comparison of hormonal production before GnRH injection demonstrated significantly lower basal T levels in B-thalassemia patients compared to healthy controls (139.7 +/- 29.8 ng/mL vs. 540 +/- 86.9 ng/mL). A similar and significant trend was observed for peak T levels post injection (289.8 +/- 49.4 ng/mL vs. 909 +/- 185.3 ng/mL). Interestingly, however, a within-group analysis of delayed puberty and normal puberty B-thalassemia patients revealed a gradient of basal and peak T levels increasing from delayed puberty patients to normal puberty patients to health control individuals (Basal: 37.2 +/- 19.9 ng/mL → 248.6 +/- 43.7 ng/mL → 540 +/- 86.9 ng/mL; Peak: 130.1 +/- 51.3 ng/mL → 449 +/- 58.5 ng/mL → 909 +/- 185.3 ng/mL).

Given the increases in bone density, muscle mass and body fat that are engendered by T production during puberty, diminished T levels in these patients directly contribute to decreased exercise capacity [11]. One consequence of this delay is elevated risk of fracture due to reduced bone mineral density [11]. Moreover, smaller physical body size has been shown to correlate with reduced lung volume, which directly limits oxygen and carbon monoxide diffusion in the alveoli [20]. Taken together, these data describe the extent to which pubertal delay can disrupt exercise capacity in TM patients.

EXERCISE MODEL

Cumulatively considering the range of symptoms and severities experienced by TM patients, I propose a framework through which clinicians can provide relevant and safe exercise recommendations. Specifically, this model makes recommendations by evaluating symptoms in terms of risk as it relates to exercise duration, frequency, and intensity.

Patients with cardiac iron burden

Excess iron levels burden the cardiovascular system, with thinner heart walls arising due to oxidative damage in cells [5]. Patients with substantial cardiac iron burden should be advised to avoid overexertion, as efforts that cause shortness of breath could be a sign of left ventricular diastolic dysfunction, one risk factor of heart attack [6]. TM patients present with a wide range of cardiac iron levels, and cardiovascular damage varies accordingly. While Sohn et al.'s primary analysis [5] focuses on the relationship between exercise capacity and cardiac iron levels as determined by MRI, cardiac MRI is an expensive and impractical method for general use in a clinical setting. Despite some conflicting concerns about their accuracy, blood ferritin measurements have been described as a widely available clinical alternative for assessing cardiac iron burden when cardiac MRI is unavailable [22].

Previous studies describe ferritin levels >500 ng/mL as elevated and levels >1000 ng/mL as abnormally high [23] [Table 2]. In patients with abnormally elevated serum ferritin levels, cardiac iron burden can be considered substantial. As such, these patients should be advised to strictly perform low intensity exercise. High and moderate intensity exercise should be avoided due to risk of heart attack. With regards to frequency, patients should be allowed to exercise moderately provided sufficient recovery time. Similarly, moderate durations can be permitted as well given the low risk of cardiovascular dysfunction and potential for cardiovascular benefit [18].

Table 2. Range of serum ferritin levels categorized from low to abnormally high in TM patients [14]. All values are in ng/mL.

< 20	20–500	500–1000	> 1000 ng/mL
Low	Normal	Elevated	Abnormally high

Table 3. Range of Hb levels categorized from low to abnormally high in TM patients [10][23].

Hb < 12 g/dL	Hb > 12 g/dL	Hb > 17.5
"LowAnemic"	"High Anemic" (Normal)	Abnormally High (in males)

Table 4. Range of Basal T levels for use when assessing pubertal state in TM patients [7].

< 200 ng/mL	200–450 ng/mL	> 450 ng/mL
Delayed Puberty	Likely Delayed Puberty	No pubertal delay

Exercise Recommendations for Patients with Thalassemia Major

Patients with anemia

Patients with anemia are likely to experience fatigue and nausea while exercising due to inhibited oxygen carrying capacity. In light of their lower VO_2 max and oxygen pulse levels, severely anemic TM patients face inherent difficulty when performing all types of exercise [3][5]. However, this does not mean that they should avoid exercise altogether. Similar to other TM symptoms, anemia severity ranges widely in a clinical setting. Patients who are deemed capable of low and moderate forms of exercise can be advised accordingly.

While VO_2 max and oxygen pulse provide a sound basis for assessing exercise capacity, these measurements are not commonly available in a clinical setting. In contrast, Hb levels are easily obtained from serum samples and—given their strong predictive value for VO_2 max and oxygen pulse—provide a suitable alternative to assess anemia severity [5]. This model categorizes anemic severity as low, high, and abnormally high based on Hb values shown in Table 3. Generally, patients with lower Hb levels should correspond to lower intensity, lower frequency, and lower duration exercise recommendations.

Patients with severe anemia will quickly shift into an anaerobic state, resulting in increased lactic acid production. As lactic acid accumulates throughout the body, it alters the physical structure of RBCs. This deformity can increase rates of hemolysis, exacerbating other TM symptoms and further contributing to feelings of exhaustion and fatigue [15][25][26].

Thus, low intensity exercise is advised for patients with high and abnormally high Hb levels. In TM individuals with low Hb levels, intensity recommendations should be made on a case-by-case basis while considering other symptoms and prioritizing patient safety. With regards to frequency, moderate exercise can be permitted with sufficient time for lactic acid levels to decrease after each bout of exercise. Duration can be low or moderate as long as patients avoid dehydration and maintain submaximal intensity.

Patients with pubertal delay

The primary consequences of pubertal delay in the context of exercise stem from diminished height, muscle mass, fat mass, and bone mineral density [11]. While healthy individuals experience T-related increases in myofibrillar protein synthesis, total bone matrix and calcium retention that exert protective effects on the body, developmental delays in TM patients place them at greater risk of injury [27]. In particular, fracture can occur during weight-bearing activities or other forms of skeletal stress [12].

Similar to previous studies, this model employs serum testosterone levels in order to assess a patient's pubertal development [8]. Pubertal status can be classified in TM patients as delayed, likely delayed or not delayed (Table 4). As is the case with other symptoms, exercise recommendations should correspond to the extent of pubertal delay. If a TM patient presents with no delay, for example, recommendations may be less restrictive. In delayed patients, however, high intensity exercise—especially weight-bearing exercise, should be avoided. With regards to frequency and duration, recommendations can be less restrictive. As long as patients exercise submaximally and suffer from no disruptive injuries, moderate frequency and duration exercise should not place individuals at risk.

CONCLUSIONS

TM patients should be advised to avoid high intensity exercise regardless of their symptom severity. When cardiac iron burden is elevated, clinicians can encourage low intensity exercise at moderate frequency and duration. Similarly, patients with mild to severe anemia should be advised to perform bouts of low intensity exercise with moderate frequency and duration.

Finally, patients with delays or likely delay in puberty are recommended to partake in low intensity exercise of short duration with moderate frequency.

While the framework proposed here fills a gap in the literature by providing a basis for clinical recommendations, several limitations must be acknowledged. First, this model considers the severity of individual symptoms as they relate to exercise capacity. Clinically, however, patients almost always present with multiple symptoms present. This is especially true with TM patients given the closely pathophysiological connection among symptoms. Nevertheless, this model provides a foundation from which clinicians can consider symptom severity and provide recommendations.

Another limitation was the use of ferritin levels and HB levels to assess the extent of cardiac iron burden and anemia, respectively. While these measurements are common enough in a clinical setting to be of practical use, other measurements such as VO_2 max, cardiac R2, and bone mineral density can provide more accurate assessments when available.

Overall, clinicians should prioritize safety when advising patients on exercise recommendations. If they are able to do so, regular exercise can provide patients with one tool to improve their physical and emotional well-being.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Ellen Fung and Michael Amano for their help and guidance in conducting research and proofreading this paper.

REFERENCES

[1] Muhammed Shahid, Saima Sadaf, and Hira. "Sequence Analysis of Human HBA and HBB Thalassemia Genes." *Sequence Analysis of Human HBA and HBB Thalassemia Genes - IJSER Journal Publication*, Volume 6, Issue 8, Aug. 2015

[2] Marie, Eman Abdulla. "THE CLINICAL PROFILES AND THE HEMATOLOGIC CHARACTERISTICS OF THALASSEMIA AMONG PATIENTS AT IBN-ALATHEER HOSPITAL IN MOSUL

CITY." *International Journal of Current Research*, 31 Mar. 2019.

[3] Kuhn V, Dierdich L, Stevenson Keller TC, Kramer CM, et al. "Red Blood Cell Function and Dysfunction: Redox Regulation, Nitric Oxide Metabolism, Anemia." *Antioxidants & Redox Signaling*, 1 May 2017.

[4] Pivkin, Igor V, Peng Z, Karniadakis GE, et al. "Biomechanics of Red Blood Cells in Human Spleen and Consequences for Physiology and Disease." *Proceedings of the National Academy of Sciences of the United States of America*, 12 July 2016.

[5] Sohn, Eugene Y, Kato R, Noetzli LJ, et al. "Exercise Performance in Thalassemia Major: Correlation with Cardiac Iron Burden." *American Journal of Hematology*, Mar. 2013.

[6] Gujja, Pradeep, Rosing Douglas R, Tripodi Dorothy J, et al. "Iron Overload Cardiomyopathy: Better Understanding of an Increasing Disorder." *Journal of the American College of Cardiology*, 21 Sept. 2010.

[7] Jeney, Viktória. "Clinical Impact and Cellular Mechanisms of Iron Overload-Associated Bone Loss." *Frontiers in pharmacology*, vol. 8, 77. 21 Feb. 2017.

[8] Al-Rimawi, Hala Saleh, Jallad MF, et al. "Pubertal Evaluation of Adolescent Boys with Beta-Thalassemia Major and Delayed Puberty." *Fertility and Sterility*, Oct. 2006.

[9] Chulani, Veenod L, and Lonna P Gordon. "Adolescent Growth and Development." *Primary Care*, Sept. 2014.

[10] Siervogel R, M, Demerath E, W, Schubert C, et al. "Puberty and Body Composition." *Horm Res*, vol. 60, 2003, 36-45.

[11] Fung, Ellen B, Vichinsky Elliott P, Kwiatkowski Janet L, et al. "Characterization of low bone mass in young patients with thalassemia by DXA, pQCT and markers of bone turnover." *Bone*, vol. 48, 6, 2011, 1305-12.

[12] Kam, D. de, Smulders E, Weerdesteyn V, Smits-Engelsman BC. "Exercise Interventions to Reduce Fall-Related Fractures and Their Risk Factors in Individuals with Low Bone Density: a Systematic Review of Randomized Controlled Trials." *SpringerLink*, 7 May 2009.

[13] Sharma A, Madaan V, Petty FD. "Exercise for Mental Health." *Prim Care Companion J Clin Psychiatry*. 8(2):106, 2006.

[14] Segal, R, C. Zwaal, E. Green, et al. "Exercise for people with cancer: a systematic review." *Current oncology (Toronto, Ont.)*, vol. 24, 4, 2017, e290-e315.

[15] Connes, Philippe & Machado, Roberto & Hue, et al. "Exercise limitation, exercise testing and exercise recommendations in sickle cell anemia." *Clinical hemorheology and microcirculation*. 49, 2011, 151-63.

[16] Wasfy, Meagan M, and Aaron L Baggish. "Exercise Dose in Clinical Practice." *Circulation*, 7 June 2016.

[17] Papassotiriou, I., Traeger-Synodinos J, Promé D, Kister J, et al. "Association of Unstable Hemoglobin Variants and Heterozygous β-Thalassemia: Example of a New Variant Hb Acharnes or [β53(D4) Ala → Thr]." *Wiley Online Library*, 27 Oct. 1999.

[18] Nystoriak, Matthew A, and Aruni Bhatnagar. "Cardiovascular Effects and Benefits of Exercise." *Frontiers in cardiovascular medicine*, vol. 5, 135, 28 Sep. 2018.

[19] Woodson, R. D., Wills RE, Lenfant C. "Effect of Acute and Established Anemia on O2 Transport at Rest, Submaximal and Maximal Work." *Journal of Applied Physiology*, 1 Jan. 1978.

[20] Beers, Eduard JP, Mart N, et al. "Exercise Tolerance, Lung Function Abnormalities, Anemia, and Cardiothoracic Ratio in Sickle Cell Patients." *Wiley Online Library*, 16 May 2014.

[21] Stagi, Stefano, Cavalli L, Iurato C, et al. "Bone metabolism in children and adolescents: main characteristics of the determinants of peak bone mass." *Clinical cases in mineral and bone metabolism : the official journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases*, vol. 10, 3, 2013, 172-9.

[22] Ajlan, Amr, Vipraskit Vip, Aydinok Yesim, et al. "MRI for the Diagnosis of Cardiac and Liver Iron Overload in Patients with Transfusion-Dependent Thalassemia: An Algorithm to Guide Clinical Use When Availability Is Limited." *Wiley Online Library*, 15 Mar. 2018.

[23] "Ferritin Test." *Mayo Clinic*, Mayo Foundation for Medical Education and Research, 10 Feb. 2017,

[24] "High Hemoglobin Count." *Mayo Foundation for Medical Education and Research*, 25 Jan. 2019.

[25] Smith, J A, Telford RD, Kolbuch-Braddon M, et al. "Lactate/H+ Uptake by Red Blood Cells during Exercise Alters Their Physical Properties." *European Journal of Applied Physiology and Occupational Physiology*, 1997.

[26] Connes, Philippe, Bouix Didier, Py Guillaume, et al. "Opposite Effects of in Vitro Lactate on Erythrocyte Deformability in Athletes and Untrained Subjects." *Clinical Hemorheology and Microcirculation*, 1 Jan. 2004.

[27] Kraemer WJ and Rogol AD. "Resistance Exercise and Testosterone." *The Endocrine System in Sports and Exercise*. 23, 2005.

The Effect of Rising Global Temperatures on Methanogenic Archaea in the Arctic Permafrost

Author: Yuki Agarwala
Supervisors: Professor Kobayashi, Mr. Imajo, Mr. Hirano
K. International School Tokyo, Tokyo, Japan
Email: agarwalayuki@hotmail.com

***Abstract* – Greenhouse gases present in the atmosphere are rising primarily due to anthropogenic reasons. However, research has shown that there is another unseen contributor, hidden in the permafrost – methanogens. These archaea produce methane from a variety of carbon compounds via methanogenesis. Since methane is more potent than carbon dioxide, it is crucial to determine the archaeal production of methane at increasing temperatures to assess the severity of the situation. In this investigation, *Methanococcus maripaludis* was inoculated at various temperatures and their growth, as well as the volume of carbon dioxide and methane, was measured using spectrophotometry and gas chromatography. The results show increased growth and methanogenesis between 10 ℃ and 15 ℃. This has significant implications on the severity of global warming because once the permafrost reaches this temperature, the methane contributions of the methanogens might accelerate the phenomenon unexpectedly, beyond control. However, archaeal growth could slow due to many factors including the lack of substrates and further investigation is necessary for the application of the findings in the real-world context.**

Key Words – Archaea, greenhouse effect, gas chromatography, methanogenesis, methanogens, permafrost, spectrophotometry

INTRODUCTION

I. Research in a Global Context

Greenhouse gases including methane and carbon dioxide absorb infrared radiation from the sun as well as radiation that is reflected back from the ground, resulting in further trapping of heat in the Earth's atmosphere. This increases global temperatures in a phenomenon known as global warming which has severe consequences on the environment, including drastically changing climates, extreme weather patterns, and the rapid melting of sea ice and permafrost [1][2]. While we are responsible for elevated greenhouse gas levels in the atmosphere, it is crucial to recognize that we may not be the only culprit.

Methanogens, or anaerobic archaeal psychrophiles, convert various carbon compounds to methane via a process known as methanogenesis [3]. At cold temperatures, methanogenic metabolism is slowed due to reduced enzyme activity, but at higher temperatures, methanogenic archaea actively metabolize, producing methane which is released through the ice. Methane is a greenhouse gas 30 times more potent than carbon dioxide at trapping heat. The heat then melts the ice, leading to a vicious cycle of warming and melting of ice [4].

II. Justification of M. maripaludis and Research Method

There are three types of methanogenesis: hydrogenotrophic methanogenesis which requires hydrogen and carbon dioxide, acetoclastic methanogenesis or the synthesis of methane from acetate, and methylotrophic methanogenesis which is the utilization of a wide range of methylated compounds [5][6]. Since most methanogens are easily affected by sub-optimal conditions, a hydrogenotrophic methanogen, *Methanococcus maripaludis*, was chosen as it has a simple metabolic pathway and can be easily cultured.

Like any other archaea, methanogens contain a multitude of different enzymes [7]. Enzymes work best at their optimal temperature and pH but can denature at extreme conditions. *M. maripaludis* grows at a moderate temperature of 20℃ and thrives in anaerobic conditions [8]. They live in an environment of 1 to 2℃ [9]. A temperature range from 5℃ to 30℃ was chosen to represent the current environment that they live in and their optimal temperature.

III. Research Statement

To investigate the consequences of rising global temperatures on methanogenic archaea in the Arctic permafrost by modelling the growth of *Methanococcus maripaludis* cultures at incremental temperatures, measuring their optical density by spectrophotometry and methane production using gas chromatography.

IV. Hypothesis for Archaeal Growth

Archaea, just like bacteria, grow in four different phases: Lag phase, log phase, stationary phase, and death phase (Figure 1). Since it was expected that the archaea would grow the fastest at their optimal temperature, it was hypothesized that at 20℃ to 25℃, the stationary phase would be reached quicker because of a higher growth rate.

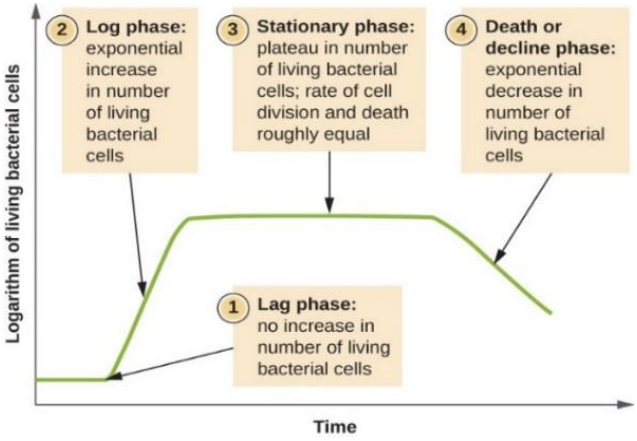


FIGURE 1. GRAPH SHOWING DIFFERENT STAGES OF BACTERIAL CULTURE GROWTH OVER TIME [10]

DEVELOPMENT OF METHOD

The experiment was conducted at the Tokyo University of Marine Sciences and Technology (TUMSAT) with Doctoral Student Imajo, and Graduate Student Hirano under the supervision of Professor Kobayashi.

Research showed that methanogenic populations could take days to peak so it was decided that the archaea would be incubated at various temperatures for 70 days [11]. Based on this, it was determined that the optical density (OD) of the archaea would be measured by a spectrophotometer (UV mini 1240, Shimadzu) to determine growth, and gas chromatography (GC-8A, Shimadzu) to measure the volume of methane gas.

A spectrophotometer passes light through a 1 ml culture sample. 460 nm was chosen for the wavelength of light because it reflects the numbers F430, an important coenzyme involved in methanogenesis, and can, therefore, be used as a measure of growth [12][13]. In the gas chromatograph (GC), hydrogen was used as the carrier gas and the sample of gas from the archaeal vials was injected into the columns. The gas compounds eluted from the column over time were recorded using Chromato-Pro Version 5.0 until it was terminated.

MATERIALS AND METHOD

A culture medium was used to cultivate the archaea to imitate the conditions of the archaea's natural environment. Medium 911 was used for *M. maripaludis*, as recommended by Dr. Mori from the National Institute of Technology and Evaluation (NITE) Biological Resource Center [14]. Three samples of *M. maripaludis* and three controls were measured at six different temperatures. A total of 36 vials of 10 ml of culture medium were required, but 40 vials (400 ml) were prepared to allow for errors. *M. maripaludis* was purchased from NITE Biological Resource Center.

I. Preparation of the culture medium

Medium 911 was required to culture *M. maripaludis*. A beaker with 350 ml of distilled water was prepared and placed on a stirrer. 0.8 ml of a trace elements solution was added, as stated on the NBRC website [14]. Further, 0.4g of resazurin was added because it is an indicator of anaerobic conditions, which is required for culturing *M. maripaludis* (Figure 2). 1.22g of hydrated magnesium chloride, 0.06g of hydrated calcium chloride, 0.216g of ammonium chloride, 0.12g of sodium sulfate, 10g of sodium chloride, 0.008g of bacterial yeast, and 0.32g of sodium acetate were added to the same solution and mixed thoroughly at room temperature. The solution was then poured into a 500 ml measuring cylinder and distilled water was added to make the volume 400 ml.

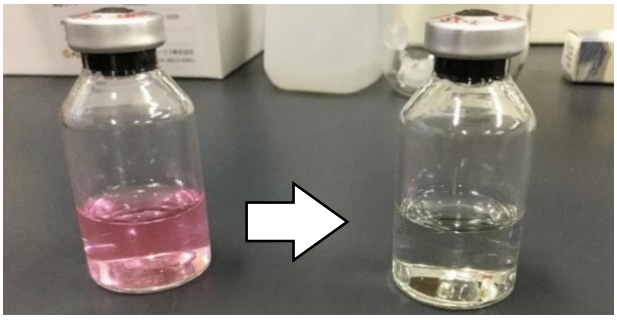


FIGURE 2. RESAZURIN ADDED TO CULTURE MEDIUM WITH OXYGEN (LEFT) AND WITHOUT OXYGEN (RIGHT)

II. Preparation of the Vials

In total, 40 vials were prepared to culture bacteria, with each vial having a volume of 20 ml. A pipette was used to transfer 10 ml of the culture medium into the 20 ml vials. A cap was screwed on top of each of the 40 vials. A gas exchange machine was used to bubble hydrogen and carbon dioxide gas through the solution (Figure 3). They were then autoclaved at 120℃ and 150 kPa for 15 minutes. A syringe was used to add 0.015g of sodium carbonate, 0.005g of hydrated sodium sulfide, 0.005g of cysteine-HCl, 0.8 ml of the KP7 buffer, and 0.8 ml of the vitamin solution to each vial. Hydrogen and carbon dioxide gas were bubbled through the culture medium once again, to remove any oxygen that could have entered. This was done by inserting another needle in the vial to allow for the passing of gas via a gas exchange machine (GR-8, Sanshin) (Figure 3).

III. Addition of Archaea

1 ml of *M. maripaludis* was transferred to 18 of the vials in a clean bench. The vials underwent gas displacement again to displace any other gases which could have entered while the archaeal solutions were added. It was made sure that the vials were colorless by this stage, a confirmation of anaerobic conditions (Figure 2). Three vials without archaea (control) and three vials with them were incubated at 5℃, 10℃, 15℃, 20℃, 25℃, and 30℃.

FIGURE 3. GAS EXCHANGE MACHINE DISPLACING OXYGEN GAS WITH CARBON DIOXIDE AND HYDROGEN GAS

IV. Measuring the Optical Density and Gas Chromatography

On the 10th, 20th, 30th, and 70th day from inoculation, the OD of the archaea was measured, and a sample viewed under an optical microscope (BX51, Olympus) with 40x magnification. The OD was measured at a wavelength of 460 nm, first by measuring the amplitude of the culture medium and setting it to 0 as a reference. After this, 1 ml of the archaeal solution was measured, and its OD was recorded. The GC was set to 250℃, 31.53 psi and the hydrogen gas flow was started at 30 ml/min. The gas composition was investigated for 0.05 cm^3 of gas.

Note that only one trial was conducted for OD and GC due to the complexity of the experiment. The decision was made in a discussion with my supervisor.

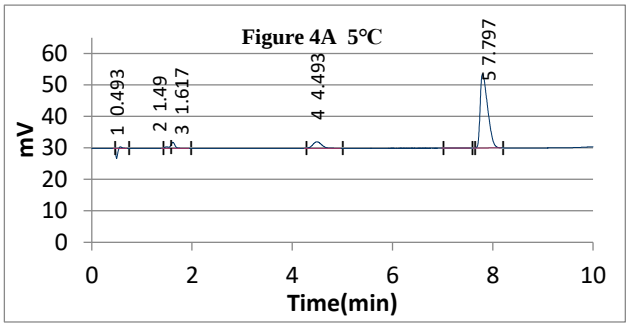
RESULTS

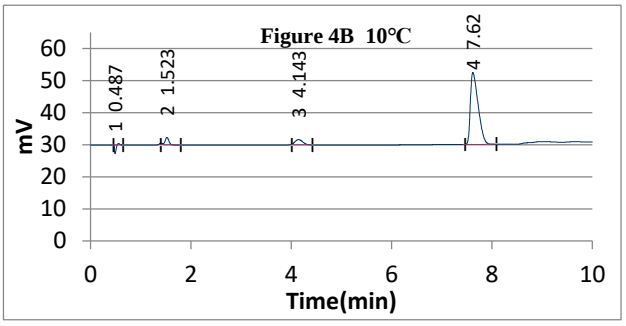
Qualitative and quantitative data were taken for the inoculations at the various temperatures. OD and GC were the quantitative measurements and microscopic pictures were taken as the qualitative data.

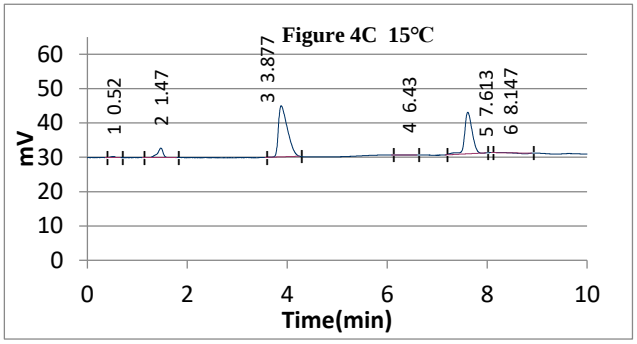
I. Quantitative Results

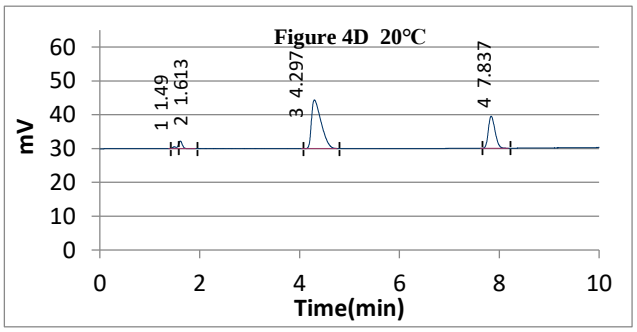
For gas chromatography, the instrument that was used recorded a peak at 3-4 minutes for methane and 8-9 minutes for carbon dioxide according to its instruction manual. The area under the peak for the gases was then calculated using Chromato-PRO version 5.0.

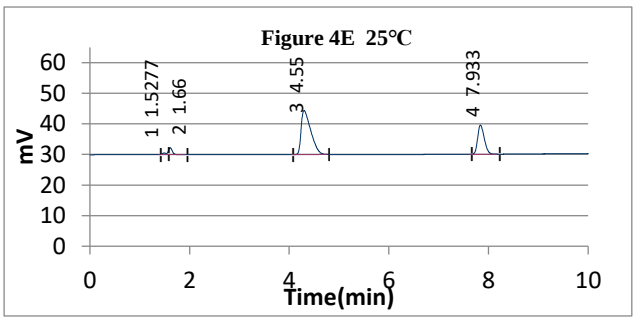
Due to the time limitations and complexities of the experiment, GC was only performed at 30 days. 30 days was chosen in consultation with my supervisor because it was assumed to be past the period of maximum growth for the archaea, resulting in the highest possible measurements. Figures 4A to 4F show gas chromatograms for *M. maripaludis* at 30 days at all temperatures.

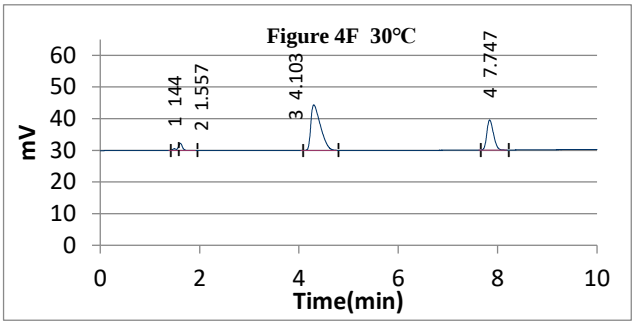








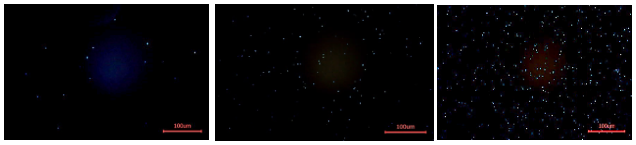




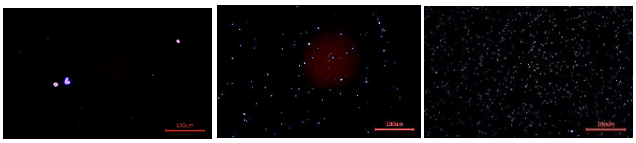
FIGURES 4A – 4F GAS CHROMATOGRAMS FOR *M. MARIPALUDIS* AT 30 DAYS AT ALL TEMPERATURES.

II. Qualitative Results

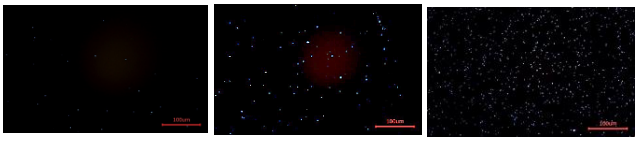
Figures 5-8 show the images taken by an optical microscope at 40x magnification. Due to a large number of images, only a few temperatures for growth at different days are represented.



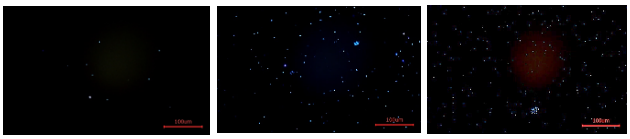
FIGURES 5A TO 5C (LEFT TO RIGHT)
GROWTH OF *M. MARIPALUDIS* AT 10 DAYS AT 5° C, 15°C, 25°C



FIGURES 6A TO 6C (LEFT TO RIGHT)
GROWTH OF *M. MARIPALUDIS* AT 20 DAYS AT 5° C, 15°C, 25°C



FIGURES 7A TO 7C (LEFT TO RIGHT)
GROWTH OF *M. MARIPALUDIS* AT 30 DAYS AT 5° C, 15°C, 25°C



FIGURES 8A TO 8C (LEFT TO RIGHT)
GROWTH OF *M. MARIPALUDIS* AT 70 DAYS AT 5° C, 15°C, 25°C

III. Processed Results

Optical density values range from 0 to 1, with 1 representing a high concentration of archaea with full absorbance of light and 0 representing no archaea and transparency. Figure 9 represents the processed OD values.

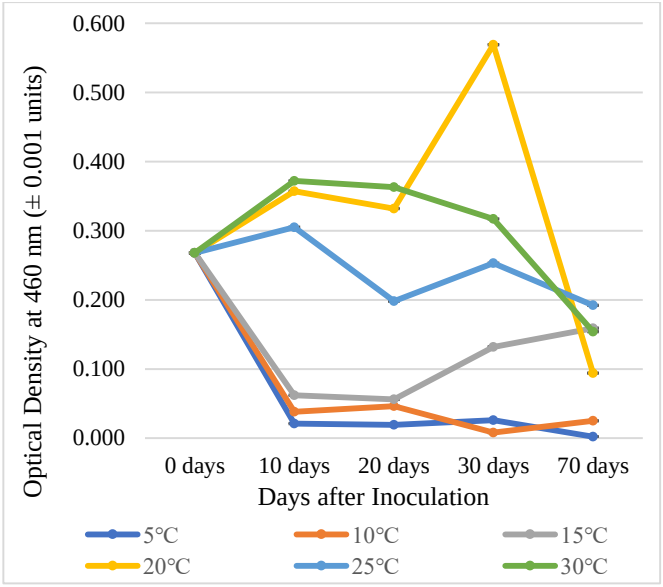


FIGURE 9. OPTICAL DENSITY OF *M. MARIPALUDIS* AT 5°C, 10°C, 15°C, 20°C, 25°C AND 30°C AT 10 DAYS, 20 DAYS, 30 DAYS AND 70 DAYS (ERROR BARS FOR OPTICAL DENSITY: ±0.001 UNITS)

ANALYSIS & DISCUSSION

The optical density measurements in Figure 9 show archaeal growth with the data at most temperatures following the ideal curve as shown in figure 1. The archaea which were cultivated at a lower temperature, 5°C, 10°C and 15°C were mostly increasing even after 70 days, but archaea inoculated at higher temperatures, 20°C, 25°C, and 30°C decreased at 70 days. The sudden peak at 20°C after 30 days, however, was an anomaly. This could have been due to microscopic bubbles that were formed in the 1ml sample of the spectrophotometer. These methanogens tend to form clumps (Figure 6A), and if these are part of the sample, the OD value could come out to be higher than expected. Although the solutions were thoroughly mixed before extracting 1ml, the aggregates did not come apart, and these could be a possible cause of the anomaly. Conducting multiple trials and taking the average of the results could reduce such random errors.

In figures 4A to 4F, small deviations can be seen around 2 minutes, due to tiny amounts of air that seeped into the syringe which was used to input gas. The chromatograms show decreasing peak heights at 8-9 minutes at increasing temperatures representing increasing consumption of carbon dioxide. The figures also show higher peaks at 3-4 minutes at increasing temperatures, representing increasing methane synthesis. Therefore, the peaks on the chromatogram show accelerated methanogenesis at higher temperatures, and the peak area under the curve was calculated to determine the exact volume of gas for comparison.

The program Chromato-PRO was used to calculate the area under the curve. The volume of methane and carbon dioxide for each temperature at 30 days was plotted. The value on the vertical axis was calculated from a calibration based on 0.05 cm^3 of gas representing 8,113,340 area.

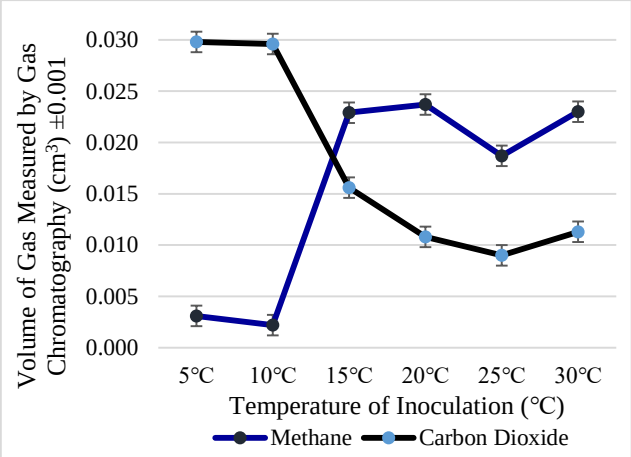


FIGURE 10. VOLUME OF CARBON DIOXIDE AND METHANE PRODUCED BY *M. MARIPALUDIS* INOCULATED AT 5°C, 10°C, 15°C, 20°C, 25°C, AND 30°C AFTER 30 DAYS AS MEASURED BY GAS CHROMATOGRAPHY (ERROR BARS FOR VOLUME OF GAS: ±0.001)

As seen in figure 10, the volume of carbon dioxide and methane had an inverse relationship. According to this result, methanogens use carbon dioxide and produce methane, intensifying global warming as methane is a much more potent greenhouse gas than carbon dioxide [4]. Although the optimal temperature of *M. maripaludis* (37°C) was not tested, the trend of the graph shows that methane production was high at higher temperatures [14]. Based on this, it was inferred that at 37°C, the growth would be even faster. The methanogens became increasingly active after 10°C as shown by the rise of methane production. Consequently, testing a 1°C range from 10°C to 15°C would give a more precise value of the specific temperature at which they become more active. The growth time was fastest at 30°C as measured by optical density, rising from 0.268 to 0.372 within the first 10 days of inoculation. This shows that as the environmental temperatures become closer to their optimal, the methanogens will grow more rapidly, producing more methane.

EVALUATION

I. Statistical Analysis

Due to the complexities of this experiment, only one trial was performed for OD and GC, which made measuring the precision and accuracy of the data difficult. Although many articles in this field were reviewed, no literature value could be found for the measurements of OD or GC for the archaea. To measure reliability, it was rationalized that for any given temperature, the volume of methane produced would always be a function of the concentration of the F430 coenzyme. Based on this, the OD and GC values should ideally give a linear function if plotted, so the Pearson correlation's coefficient was calculated. The 6 points from all the temperatures at 30 days gave a coefficient of 0.59, which is a moderate correlation. Out of this, 4 points had a linear relationship with a coefficient of 0.99. Although 0.59 indicates low reliability, the results can be considered reliable because, without the anomalies, the data forms a near-perfect line (R value = 0.99) and the anomalies can be attributed to natural variation as well as random errors. For example, measuring the OD of the archaea was difficult as they tend to aggregate (Figure 6A), and the 1ml sample of the spectrophotometer might have contained the aggregate, causing exceptionally high OD values. Although care was taken, it was difficult to ensure this, which could have affected the measurements.

II. Methodological Limitations

For a few hours every 10 days, the vials were taken from the incubators and left in the laboratory while microscope images, OD_{460}, and GC measurements (for 30 days only) were taken. For vials incubated around 5°C, the sudden change to 25°C may have affected archaeal populations. These drastic differences in temperature are not present in the methanogen's natural environment. To improve, a sample should be taken in a clean bench which is set to the temperature of the incubator.

In this experiment, light of the wavelength 460 nm was used to measure growth because it detects a coenzyme involved in methanogenesis. An extension would be to use lights of different wavelengths which can detect a variety of organic substances. A peak other than the archaea can then be compared against a database. There is always a possibility for an unknown substance to be hindering the results, and this would be one way to further investigate it.

III. Global Applications

In this experiment, temperature was the limiting factor for the growth of *M. maripaludis*. The results clearly showed that at higher temperatures, more carbon dioxide was converted into methane. Although increases in methane production due to temperature rises could occur, in reality, there are many other influential factors as well. When permafrost is exposed to warmer temperatures, it melts and releases the gases trapped in it, including hydrogen and carbon dioxide. Since these gases are used to produce methane, their unavailability would lead to decreased methanogenesis. This was not modeled in the experiment because the archaea's methane production cannot be tracked in an open vessel. In ecosystems, once the gasses escape from the permafrost, the archaea are permanently deprived of them. As the temperature increases, the permafrost could defrost other organisms, which could result in interspecies competition for substrates, reducing methanogenic populations. The melting permafrost would also cause more oxygen exposure which could potentially kill some methanogen species. Although it is shown that methanogens will actively produce methane when the temperature increases past 10°C, the balance between interspecies competition and oxygen levels may determine the impact of these archaea on global warming.

CONCLUSION

The focus of this investigation was to understand the consequences of global warming on methanogenic archaeal communities in the permafrost and vice versa. Since *M. maripaludis* showed higher metabolic activity at temperatures near its optimal, it can be concluded that as global warming proceeds and temperatures rise, more methanogens in the Arctic permafrost will reach their optimal temperature and produce consequential amounts of methane. This, in turn, will cause temperatures to rise faster. As global warming increases the temperature of the methanogens' habitat, bringing it above 10℃, methanogens undergoing more rapid methanogenesis would result in more methane will be produced. When the effects of millions of these organisms over decades are summated, there will be a significant increase in methane in the atmosphere. Since methane is known to be a much more potent greenhouse gas than carbon dioxide [4], it can be expected that the increased methanogenesis will intensify global warming and cause a positive feedback loop, until substrates are available for methanogenesis.

ACKNOWLEDGMENTS

I would like to give a heartfelt thank you to Associate Professor Takeshi Kobayashi at the Tokyo University of Marine Sciences and Technology for giving me the opportunity to conduct my experiment. I would also like to thank Mr. Takumi Imajo, and Mr. Shiori Hirano for guiding me through this rigorous experiment. I would also like to give a special thank you to Dr. Koji Mori from the NITE Biological Resource Center for providing me with the archaea and kindly introducing me to Professor Kobayashi.

REFERENCES

[1] Climate Change. Retrieved February 12, 2018, from http://www.un.org/en/sections/issues-depth/climate-change/

[2] Press, A. (2017, December 12). Arctic permafrost thawing faster than ever, US climate study finds. Retrieved January 23, 2018, from https://www.theguardian.com/environment/2017/dec/12/arctic-permafrost-sea-ice-thaw-climate-change-report

[3] Graham, D. E., Wallenstein, M. D., & Vishnivetskaya, T. A. (2012). Commentary. *Microbes in thawing permafrost: the unknown variable in the climate change equation*. Retrieved January 24, 2018.

[4] Abraham, J. (2015, October 13). Methane release from melting permafrost could trigger dangerous global warming | John Abraham. Retrieved January 23, 2018, from https://www.theguardian.com/environment/climate-consensus-97-per-cent/2015/oct/13/methane-release-from-melting-permafrost-could-trigger-dangerous-global-warming

[5] Allen, M. A., Lauro, F. M., Williams, *et al* (2009). The genome sequence of the psychrophilic archaeon, Methanococcoides burtonii: the role of genome evolution in cold adaptation. *The ISME Journal*. Retrieved February 14, 2018, from http://www.nature.com/articles/ismej200945.pdf

[6] As permafrost soils thaw, soil microbes amplify global climate change. (2014, October 22). Retrieved February 04, 2018, from https://www.sciencedaily.com/releases/2014/10/141022131434.htm

[7] Shima, S., Warkentin, E., Thauer, R. K., & Ermler, U. (n.d.). Structure and function of enzymes involved in the methanogenic pathway utilizing carbon dioxide and molecular hydrogen. Retrieved April 11, 2018, from https://www.ncbi.nlm.nih.gov/pubmed/16233244

[8] Methanococcus maripaludis. (n.d.). Retrieved July 24, 2019, from https://microbewiki.kenyon.edu/index.php/Methanococcus_maripaludis

[9] Nichols, D. S., Miller, M. R., Davies, N. W., Goodchild, A., Raftery, M., & Cavicchioli, R. (2004). Cold Adaptation in the Antarctic Archaeon Methanococcoides burtonii Involves Membrane Lipid Unsaturation. *Journal of Bacteriology, 186*(24), 8508-8515. doi:10.1128/jb.186.24.8508-8515.2004

[10] How Microbes Grow. (n.d.). Retrieved June 28, 2018, from https://courses.lumenlearning.com/microbiology/chapter/how-microbes-grow/

[11] Franzmann, P. D., Liu, Y., Balkwill, D. L., Aldrich, H. C., Marcio, E. C., & Boone, D. R. (1997). Methanogenium fngidum sp. nov., a Psychrophilic, H,-Using Methanogen from Ace Lake, Antarctica. *International Journal of Systematic Bacteriology*. Retrieved June 26, 2018, from http://www.microbiologyresearch.org/docserver/fulltext/ijsem/47/4/ijs-47-4-1068.pdf?expires=1530020069&id=id&accname=guest&checksum=5092E3CDA72188B820CC017ED1764821

[12] Buan, N. R., & Metcalf, W. W. (2010). HUG 10, ff. 020r, inc. AC conchoides. D polus. EF regula. *Molecular Biology*. doi:10.1163/2468-0303-cohu_10-026

[13] Duin, E. C. (2009). Role of Coenzyme F430 in Methanogenesis. *Springer*, 352-374. doi:10.1007/978-0-387-78518-9_23

[14] Catalogue Detail Information. (2004). Retrieved June 30, 2018, from http://www.nbrc.nite.go.jp/NBRC2/NBRCCatalogueDetailServlet?ID=NBRC&CAT=00101831&lang=jp

Comparative Analysis of IMUs and Motion Detecting Technology for Measuring Patients' Lumbar and Cervical Spines' Range of Motion

David Bass
The American School in Japan, Tokyo, Japan
Email: 21bassd@asij.ac.jp

Abstract - **Chiropractors and doctors often analyse their patients' musculoskeletal health by measuring their spinal range of motions (ROM) through estimation techniques, which do not always yield accurate results; this could potentially result in an inaccurate diagnosis of patients' conditions. In order to prevent such errors from occurring, technologies such as the Inertial Measurement Unit (IMU) and Motion Detection can be used on patients. Such implementations would allow clinics to accurately measure their patients' ROMs. This study involves the application of the IMU and Motion Detector in clinical settings in order to assess their reliability for measuring ROMs. The experiment required the participant to perform four specific movements while wearing the IMU device and simultaneously being recorded by the Motion Detector. Data from each trial was retrieved from both devices and comparatively analyzed. The results indicated that the two sets of data collected from the IMU and Motion Detector contained several similarities but mostly contained discrepancies. These discrepancies were likely caused by confounding variables but not due any defects in the technology's measuring capabilities.**

Key Words - Back Posture, Chiropractor, Inertial Measurement Unit, Kinect, Motion Detection

INTRODUCTION

I. Background

In the United States, 80% of adults are bound to experience lower back pain within their lifetime. [1] Back pain is responsible for the loss of 264 million work days annually. Additionally, back pain is correlated with depression, mobility issues, and lower quality of sleep. [2,3,4] Most cases of back pain can be treated with professional medical support. However, in order to provide the correct treatment, it is critical for a proper diagnosis to be conducted beforehand.

Chiropractors and doctors who are responsible for patient diagnosis utilize a myriad of techniques. A commonly used technique involves the assessment of a patients' spinal range of motion (ROM), which requires patients to extend, flex, and rotate regions of their vertebrae. [5] In this study, the lumbarsacral and cervical regions of the vertebrae will be evaluated. ROM can help professionals evaluate a multitude of factors of a person's musculoskeletal health. However, many chiropractors and doctors who conduct ROM tests rely on pure observation or best estimates, which are prone to error. An error in evaluating a patient's ROM can result in severe consequences such as misdiagnosis, inapt treatment, and ultimately the worsening of their back conditions. In order to minimize the possibility of running into errors, there are several technologies that can be utilized to accurately evaluate patients' ROM. In this study, two technologies: 1) IMU and 2) Motion Detection will be tested on a patient. The purpose of testing these technologies is to determine and compare their abilities to produce accurate data that can be evaluated by a medical professional who needs information on their ROM. The objective of this study is to determine the accuracy of the data produced by the IMU and motion detector when being used on a patient, which can be determined by comparatively analyzing their data sets. This experiment will test the hypothesis that in a controlled setting, the data produced by the IMU and motion detector will have statistically insignificant discrepancies. This finding will allow us to determine the degree of reliance that professional chiropractors and doctors can have for data produced by these technologies.

II. Literature Review

Numerous studies in the past have used IMU and Motion Detecting technology to analyze the musculoskeletal health of researchers' test participants. Sung Wo Kan et al. developed IMU-integrated clothing that was intended to monitor the posture of patients in their daily life for spinal disease prevention; their IMU-derived data produced angles with less than an error of less than 4°, confirmed by comparing their results to a motion capture system [6]. Other researchers such as Abyarjoo et al. also created IMU-integrated clothing that produces real-time warnings when the wearer shifts to a position of "bad posture" [7]. An additional comparative study by Simpson et al. reviewed 37 articles that involved the utilization of IMUs to assess spinal posture, concluding that IMUs are able to evaluate spinal

posture with "good accuracy" [8]. With regards to using motion detection technology to analyze spinal posture, researchers such as A.P.G. Castro et al. have used *Microsoft Kinect* as a means to evaluate patients' spinal curvatures, describing the procedure as "effective" [9] for clinical and commercial implementations. Another paper published by Jamie Shotton et al. revealed that their experiments using *Microsoft Kinect* for image-based pose estimations "accurately predicted" their subjects' body joint locations [10].

METHODOLOGY

I. Equipment

This experiment involved the use of two technologies to evaluate spinal ROM: IMUs and Motion Detection.

The Inertial Measurement Unit is a device that combines the functions of a gyroscope, magnetometer, and accelerometer. These functions include the ability to detect the IMU sensor's orientation and angular velocity in relation to Earth's magnetic and gravitational fields. [11] The IMU used in this study was the *Adafruit 9-DOF Absolute Orientation IMU Fusion Breakout - BNO055* (Figure 1).

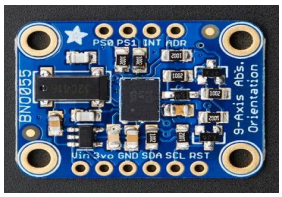


FIGURE 1
ADAFRUIT 9-DOF ABSOLUTE ORIENTATION IMU FUSION BREAKOUT - BNO055 [12]

This study uses the IMU to collect data points relating specifically to the angle of the sensor's orientation. The experiment involved placing two IMUs along the participant's back for measuring his/her lumbarsacral spine ROM, and one IMU on the Occipital Bone for measuring cervical spine ROM. The IMUs were connected to the Bluetooth Arduino Nano [13], where the coding information for data processing was located. The Arduino retrieved data from the IMUs and wirelessly sent it to a computer that was used for data collection. The wireless transmission of data to the computer was enabled by the HC-05 Bluetooth Module. This entire network of devices was transplanted onto a belt that would be strapped onto the participant (Figure 2).

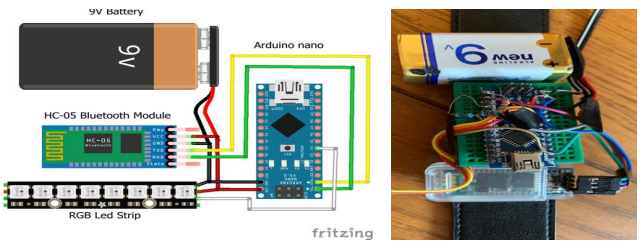


FIGURE 2
ARDUINO BLUETOOTH NETWORK [13]

The software used to collect the data was *Teraterm*, which organized the IMU data into seven columns [14]. The first column contained timestamps, the subsequent three columns contained data relating to positional orientation, and the final three columns contained data relating to angular orientation. This study focuses on the angular orientation data, in order to compare the initial and final angles of the participant's orientation while performing a ROM test.

Additionally, motion detection technology was used to track the participant's lumbosacral and cervical spines' ROM. The device used for this procedure was the *Microsoft Kinect*, which was initially a complement of the Xbox One [15]. The *Kinect* hardware consists of a camera with an infrared laser projector which enables it to analyze the three-dimensional depth of an image or video [16]. In order to utilize the *Kinect* device to perform the necessary functions for this study, the software *Kinetisense* was ordered. *Kinetisense* uses the *Kinect* device to locate a person's joints in the neck, hips, back, and other regions, and can subsequently perform a wide range of analyses based on this information [17]. This study focuses on the ROM analysis of a test subject's lumbosacral and cervical spines. The four specific functions used were: Back Flexion (bend down), Back Extension (lean back), Neck Flexion (bend head down), and Neck Extension (lean head back); in all cases, hips are stationary, legs remain straight (Table I).

II. Procedure

In order for there to be a significance in the comparison of data sets produced by the IMU and Kinetisense, the two devices must be tested simultaneously. Our test required the participant to wear the IMU device while being captured by the Kinetisense motion detector. Before each trial begins, the participant must stand upright and follow the procedure for each of the four Kinetisense functions (Back Flexion, Back Extension, Neck Flexion, Neck Extension). The procedures required in our study were based on the ROM guidelines provided by Kinetisense (Table I).

The procedures outline every standard that must be met when conducting a ROM test with the Kinetisense software. Once the participant understood the procedure, he/she was given the IMU-Arduino belt to wear. For tests that involved back extension and back flexion, two IMUs were placed on the participant's back; one on the upper thoracic region and one on the sacral region of the spine. For tests that involved neck extension and neck flexion, a single IMU was placed in the cervical region of the spine (Figure 4).

Once the Kinetisense sensor and IMU equipment is set up, the test may begin. First, the Kinetisense video must be played while the IMU devices simultaneously activate and begin recording data. When both the Kinetisense and IMU begin their forms of data collection, the participant may leave his/her initial position and perform the procedure towards the final position. Once the final position is reached, both Kinetisense and the IMU may stop data collection (Figure 3). (Note: the IMU placement on Figure 4

was placed sideways during the tests, different than the vertical placement as seen in the image).

TABLE I
PROCEDURE BASED ON KINETISENSE ROM GUIDELINES [18]

Back Flexion	Initial Position	• Stand upright • Face sideways from camera • Relax arms • Place hands in front of thighs • Plant feet together • Back straight
	Change	• Bend down • Maintain straight legs, do not bend • Arms remain relaxed and hang towards the ground
	Final Position	• Back bent down as far as possible
Back Extension	Initial Position	• Stand upright • Face sideways from camera • Relax arms on the side of your torso with palms facing back
	Change	• Extend back backwards • Keep knees complacent • Move hips forward • Push hands and arms back • Gaze at the ceiling
	Final Position	• Back extended as far back as possible
Neck Flexion	Initial Position	• Stand Upright • Face sideways from the camera • Hang arms on sides • Keep shoulders, abdomen, and torso complace
	Change	• Bend neck down • Tuck chin into chest • Do not rotate neck while bending down
	Final Position	• Neck facing ground
Neck Extension	Initial Position	• Stand upright • Face sideways from camera • Hang arms on sides • Keep shoulders and torso complace
	Change	• Tilt head back towards the ceiling • Maintain contact between upper and lower feet
	Final Position	• Neck facing ceiling

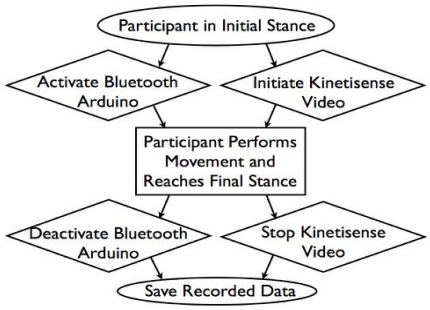


FIGURE 3
FLOWCHART FOR TESTING PROCEDURE

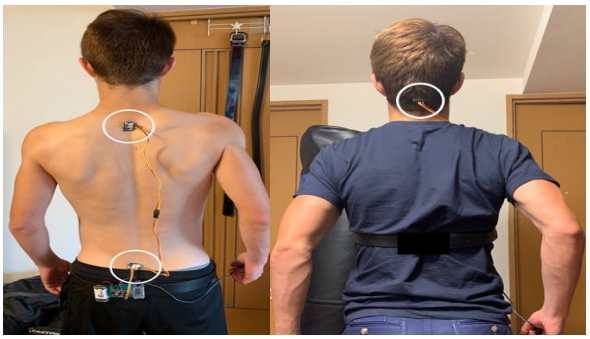


FIGURE 4
IMU PLACEMENT: BACK EXTENSION/FLEXION (LEFT), NECK EXTENSION/ FLEXION (RIGHT)

RESULTS

Each procedure (Back Flexion, Back Extension, Neck Flexion, Neck Extension) was tested four times, resulting in a total of 16 trials and 32 data values from both the IMU and Kinetisense.

Data from the IMU contained six columns regarding orientation: x, y, z, roll, pitch, and yaw. For this experiment, x, y, and z were not needed because they indicate positional orientation, not angular orientation. The subsequent columns of roll, pitch, and yaw do measure angular orientation. When the IMU is placed horizontally on the participant, roll corresponds to a twisting/spinning motion, pitch corresponds to a vertical inclination/declination, and yaw corresponds to a sideways rotation (Figure 5). In the context of this experiment, only the pitch values are relevant due to the measurement of inclination and declination, which directly correspond to what Kinetisense is also measuring — ROM of the neck and back.

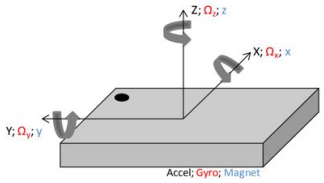


FIGURE 5
IMU: ROLL (X), PITCH (Y), AND YAW (Z) [19]

Each pitch value produced by the IMU corresponded to a timestamp. This allows a graphical representation of how the IMU angles change over time to be produced (Figure 6). To calculate the change in angle from the participant's initial and final position, the minimum and maximum values of the data set were retrieved and subtracted from each other. These values represent the ROM of the participant's lumbosacral spine and cervical spine (Table II). Sensor 2 values were not included in this study because Sensor 1 had already yielded the necessary results.

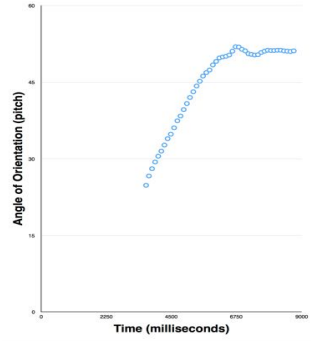


FIGURE 6
TIME VS PITCH (SENSOR 1 IMU)

Data from Kinetisense was retrieved via reports that were auto-generated by the software. For functions such as Back Flexion and Neck Extension, the reports yielded correct results. However, the Back Extension and Neck Flexion reports did not accurately correspond to the minimum and maximum angular values from the original videos; hence, the video had to be rewatched and have its values written down manually. When data from the automated report was accurate (determined by whether or not it reflects the video's results), it was noted as the final

"Δ Angle". When the reported data was inaccurate, the manually gathered results from the video were noted as the final "Δ Angle" (Table III).

TABLE II
IMU DATA: CHANGE IN ANGULAR ORIENTATION (IN DEGREES)

BACK FLEXION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle	105.6875	98.9375	116.125	121.6875
Final Angle	166.375	170.25	176	175.5625
Δ Angle	**60.6875**	**71.3125**	**59.875**	**53.875**
BACK EXTENSION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle	83.5625	81.875	86.25	77.125
Final Angle	45.1875	43.4375	44.0625	38.0625
Δ Angle	**38.375**	**38.4375**	**42.1875**	**39.0625**
NECK FLEXION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle	15.625	16.5	16.9375	14.5
Final Angle	-13.125	-9.3125	-13.75	-18.75
Δ Angle	**28.75**	**25.8125**	**30.6875**	**33.25**
NECK EXTENSION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle	23.1875	20.6875	24.8125	27.375
Final Angle	53.187	52.125	51.125	53.9375
Δ Angle	**29.9995**	**31.4375**	**26.3125**	**26.5625**

TABLE III
KINETISENSE DATA: CHANGE IN ANGULAR ORIENTATION (IN DEGREES)

BACK FLEXION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle (report)	-5.7	-6.1	-19.8	-8.1
Final Angle (report)	-67.8	-60.7	31.6	-70.3
Δ Angle	62.1	54.6	51.4	62.2
Initial Angle (video)	-5	-6	-11	-8
Final Angle (video)	-68	-61	-70	-71
Δ Angle	63	55	59	63
Actual Δ Angle	**62.1**	**54.6**	**59**	**62.2**
BACK EXTENSION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle (report)	11.4	8.2	19.3	6.2
Final Angle (report)	12.3	13.2	19.3	8
Δ Angle	0.9	5	0	1.8
Initial Angle (video)	11	9	7	6
Final Angle (video)	-27	-26	-19	-26
Δ Angle	38	35	26	32
Actual Δ Angle	**38**	**35**	**26**	**32**
NECK FLEXION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle (report)	-36.4	-36.9	-35.8	-29.2
Final Angle (report)	3.8	-39.2	-18.4	-41.5
Δ Angle	40.2	2.3	17.4	12.3
Initial Angle (video)	36	37	36	29
Final Angle (video)	13	43	18	42
Δ Angle	23	6	18	13
Actual Δ Angle	**23**	**6**	**18**	**13**
NECK EXTENSION	Trial 1	Trial 2	Trial 3	Trial 4
Initial Angle (report)	33	38.4	37.5	36.7
Final Angle (report)	53.1	57.6	63	54.2
Δ Angle	20.1	19.2	25.5	17.5
Initial Angle (video)	33	38	38	36
Final Angle (video)	53	57	63	54
Δ Angle	20	19	25	18
Actual Δ Angle	**20.1**	**19.2**	**25.5**	**17.5**

COMPARATIVE ANALYSIS

The "Δ Angle" data sets produced by the IMU and Kinetisense were compared (Table IV). Of the four functions, data collected from the Back Flexion tests produced the most similar results. Back Extension produced the second most similar results; Neck Extension third, and Neck Flexion fourth. The degree of similarity between data sets was determined by calculating the difference between the IMU and Kinetisense values, and then classifying the difference values into four categories: Exact (0.0˚~0.9˚), High (1.0˚ ~ 2.9˚), Medium (3.0˚ ~ 5.9˚), Low (5.0 ~ 9.9˚), and None (10.0˚+); in total, six trials were classified under "None", four trials under "Low", two trials under "Medium", zero trials under "High", and four trials under "Exact" (Table V). Overall, the trials revealed that the data sets produced by the IMU and Kinetisense were more often different than similar.

TABLE IV
COMPARISON OF IMU AND KINETISENSE DATA (IN DEGREES)

Back Flexion	Δ Angle		Data Discrepancy	Data Similarity
Trial #	IMU	Kinetisense		
1	60.6875	61.25	0.5625	Exact
2	71.3125	54.6	16.7125	None
3	59.875	59	0.875	Exact
4	53.875	62.2	8.325	Low

Back Extension	Δ Angle		Data Discrepancy	Data Similarity
Trial #	IMU	Kinetisense		
1	38.375	38	0.375	Exact
2	38.4375	35	3.4375	Medium
3	42.1875	26	16.1875	None
4	39.0625	32	7.0625	Low

Neck Flexion	Δ Angle		Data Discrepancy	Data Similarity
Trial #	IMU	Kinect		
1	28.75	23	5.75	Medium
2	25.8125	6	19.8125	None
3	30.6875	18	12.6875	None
4	33.25	13	20.25	None

Neck Extension	Δ Angle		Data Discrepancy	Data Similarity
Trial #	IMU	Kinetisense		
1	29.9995	20.1	9.8995	Low
2	31.4375	19.2	12.2375	None
3	26.3125	25.5	0.8125	Exact
4	26.5625	17.5	9.0625	Low

TABLE V
CLASSIFICATION OF ANGULAR DISCREPANCIES

0.0˚ ~ 0.9˚	Exact
1.0˚ ~ 2.9˚	High
3.0˚ ~ 5.9˚	Medium
6.0˚ ~ 9.9˚	Low
10.0˚ +	None

DISCUSSION AND LIMITATIONS

The frequent discrepancies between the IMU and Kinetisense results indicate that the two devices produce different values when calculating the ROM of test subjects. There may be two reasons for the discrepancies: 1) One or both of the devices contained a flaw when measuring the participants' ROM or 2) The two devices do not test for the same variables. In this study, a combination of both of these reasons may be responsible for the discrepancies produced by the IMU and Kinetisense data.

A possible flaw involving the IMU may have involved the placement of the IMU on the subject's spine and cranium. Since tape, string, and other makeshift materials were used to attach the IMU to the subject, the stability of the IMU may have deteriorated after repeated trials. However, the IMU's pure ability to measure angular velocities accurately is most likely not a reason that contributed to a possible flaw.

On the other hand, Kinetisense contained flaws that were measurable (Table VI). A "Yes" answer to "Significance of Discrepancy" indicates that the the report generated inaccurate data that differed greatly from the original video; a "No" answer indicates that the reports were accurate. None of the Back Extension data and only half of the Neck Flexion data generated by the data reports had accurately reflected data from the original video. This is likely due to bugs in the software. As a result, data had to be

manually retrieved from the video (Figure 7). In some cases, the video itself did not properly track the angles of the participant's orientation, which made it improbable to harvest accurate results from the trial. An additional factor that may have contributed to the data discrepancies is the actual angles being measured. The IMU measures angles based on the Earth's gravitational field, which uses the angle perpendicular to the ground as a reference angle. This means that all angular values produced by the IMU are based off of a 90˚ reference angle. However, Kinetisense's angular values are not based off of a 90˚ reference angle, but instead generates an angle at the joint between the neck/spinal vectors or the spinal/hip vectors of a patient (Figure 8). Since the IMU and Kinetisense measured for different angles, their different reference angles may have contributed to the discrepancies between the two sets of data.

TABLE VI
DISCREPANCY IN DATA: AUTO-GENERATED REPORT VS ORIGINAL VIDEO

Back Flexion	Trial 1	Trial 2	Trial 3	Trial 4
Discrepancy in Δ Angle	0.9	0.4	7.6	0.8
Significance of Discrepancy (y/n)	No	No	Yes	No
Back Extension	Trial 1	Trial 2	Trial 3	Trial 4
Discrepancy in Δ Angle	37.1	30	26	30.2
Significance of Discrepancy (y/n)	Yes	Yes	Yes	Yes
Neck Flexion	Trial 1	Trial 2	Trial 3	Trial 4
Discrepancy in Δ Angle	17.2	3.7	0.6	0.7
Significance of Discrepancy (y/n)	Yes	Yes	No	No
Neck Extension	Trial 1	Trial 2	Trial 3	Trial 4
Discrepancy in Δ Angle	0.1	0.2	0.5	0.5
Significance of Discrepancy (y/n)	No	No	No	No

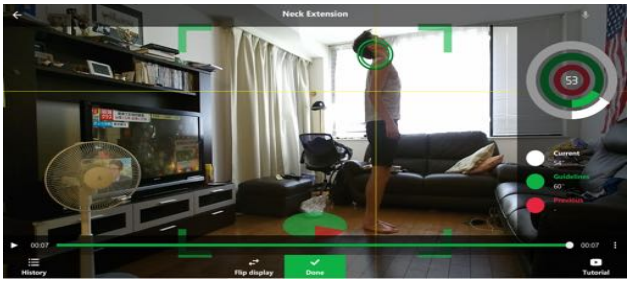


FIGURE 7
ORIGINAL VIDEO FOR NECK EXTENSION; INCLUDES ANGLE MEASUREMENT ON UPPER RIGHT SIDE OF THE INTERFACE

FIGURE 8
LEFT TO RIGHT: BACK FLEXION, NECK EXTENSION, NECK FLEXION

While discrepancies existed between two sets of data, the results in this study are valuable for a multitude of reasons. Chiropractors and doctors can examine the data sets from the IMU and Kinect Sensor to consider which technology is best tailored to their clinical needs. Chiropractors who want to use Motion Detection technology can examine the accuracy and type of data produced by Kinetisense in this study to decide what type of Motion Detecting device they would want to use for their assessments: Kinetisense, or other softwares.

In addition to data, other factors of the IMU and Kinetisense may affect a Chiropractor's evaluation of these technologies; these include: price, convenience, and the interface of the softwares. In terms of price, all of the devices involved for the setup of the IMU combined was much cheaper than the cost of the Kinetisense software license and Kinect motion detecting camera. In terms of convenience, Kinetisense is much easier to extract data from, because the data values needed are produced immediately after each trial, while the IMUs require a tedious process of data collection, interpretation and analysis. Additionally, Kinetisense does not require the attachment of devices onto patients. In terms of their interfaces, Kinetisense is much easier to navigate because the software is specifically tailored to the needs and abilities of a chiropractor or regular user. Meanwhile, the utilisation of IMUs for musculoskeletal evaluations require the combination of a myriad of softwares and equipment that may be unfamiliar to most users. Hence, in a chiropractic setting, Kinetisense would be much easier of a technology to employ. However, if a user-friendly, finished product was created using IMU technology, its level of convenience would then equal that of Kinetisense, and would be the most ideal for a chiropractic setting due to already having more accurate data as shown in this study.

In future studies, the confounding variables above can be eliminated by choosing a different motion-detection software with fewer bugs, and using standard equipment to place the IMUs onto the participant(s). Additionally, a third device such as a motion capture system with fiducial markers should be used as an accurate baseline to compare the data sets to; this will allow the experimenter to analyze which device is more reliable.

CONCLUSION

This study evaluates the reliance of the IMU and Motion Detector for calculating patients' ROMs. This study is intended to assist chiropractors and doctors to determine which methodology would be more tailored to their clinical assessments. The comparative analysis in this study indicated that the data produced by the IMU and Kinetisense contained several similarities, but mostly contained discrepancies. However, the discrepancies between the data sets were not likely due to the inaccuracy of either devices, but were most likely caused by the combination of several confounding variables. Further investigation into these technologies with more controlled variables and proper equipment will allow more precise conclusions to be established regarding the reliability of these technologies for measuring back posture and ROM.

ACKNOWLEDGEMENTS

Thank you to my mentors Abbi Hamed, Neil Nguyen, Aneesh Chand, and Tyler Kusunoki. Additionally, thank you

to the Kinetisense staff David Schnare and Codi Loran. Thank you my parents. Finally, thank you to chiropractors Dr. David Brisson from Tokyo Paramedi and Dr. Hino from Ligare Chiropractic.

REFERENCES

1. Low Back Pain Fact Sheet | National Institute of Neurological Disorders and Stroke. (n.d.). Retrieved October 1, 2019, from https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Low-Back-Pain-Fact-Sheet

2. Robertson, D., Kumbhare, D., Nolet, P., Srbely, J., & Newton, G. (2017). Associations between low back pain and depression and somatization in a Canadian emerging adult population. The Journal of the Canadian Chiropractic Association, 61(2), 96–105. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28928493

3. Sleep and Hypnosis : A Journal of Clinical Neuroscience and Psychopathology. (n.d.). Retrieved October 1, 2019, from http://www.sleepandhypnosis.org/ing/abstract.aspx?MkID=253

4. Sherwin Nicholson, H. B. S. (n.d.). Retrieved October 1, 2019, from https://lowbackpainprogram.com/side-effects-of-back-pain/

5. Cupon, L. N., & Jahn, W. T. (2003). Current standards for measuring spinal range of motion for impairment. Journal of Chiropractic Medicine, 2(1), 8–12. https://doi.org/10.1016/S0899-3467(07)60067-6

6. Kang, S.-W., Choi, H., Park, H.-I., Choi, B.-G., Im, H., Shin, D., … Roh, J.-S. (2017). The Development of an IMU Integrated Clothes for Postural Monitoring Using Conductive Yarn and Interconnecting Technology. Sensors , 17(11). https://doi.org/10.3390/s17112560

7. Abyarjoo, F., O-Larnnithipong, N., Tangnimitchok, S., Ortega, F., & Barreto, A. (2015). PostureMonitor: Real-Time IMU Wearable Technology to Foster Poise and Health. International Conference of Design, User Experience, and Usability, 543–552. https://doi.org/10.1007/978-3-319-20889-3_50

8. Simpson, L., Maharaj, M. M., & Mobbs, R. J. (2019). The role of wearables in spinal posture analysis: a systematic review. BMC Musculoskeletal Disorders, 20(1), 55. https://doi.org/10.1186/s12891-019-2430-6

9. Castro, A. P. G., Pacheco, J. D., Lourenço, C., Queirós, S., Moreira, A. H. J., Rodrigues, N. F., & Vilaça, J. L. (2017). Evaluation of spinal posture using Microsoft KinectTM: A preliminary case-study with 98 volunteers. Porto Biomedical Journal, 2(1), 18–22. https://doi.org/10.1016/j.pbj.2016.11.004

10. Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., … Blake, A. (2013). Efficient human pose estimation from single depth images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(12), 2821–2840. https://doi.org/10.1109/TPAMI.2012.241

11. Inertial Measurement - an overview | ScienceDirect Topics. (n.d.). Retrieved October 1, 2019, from https://www.sciencedirect.com/topics/engineering/inertial-measurement

12. Edge Lit Acrylic LED Sign Using Arduino. (n.d.). Retrieved October 1, 2019, from Arduino Project Hub website: https://create.arduino.cc/projecthub/theSTEMpedia/edge-lit-acrylic-led-sign-using-arduino-bbfd8e

13. Adafruit Industries. (n.d.). Adafruit 9-DOF Absolute Orientation IMU Fusion Breakout - BNO055. Retrieved October 1, 2019, from https://www.adafruit.com/product/2472

14. Tera Term Project Top Page - OSDN. (2004, November 9). Retrieved October 1, 2019, from https://osdn.net/projects/ttssh2/

15. Anthony, S. (2013, August 13). Xbox One no longer requires Kinect, but it'll still come in the box, and you'll still pay $500 - ExtremeTech. Retrieved October 1, 2019, from ExtremeTech website: http://www.extremetech.com/gaming/163796-xbox-one-no-longer-requires-kinect-but-itll-still-come-in-the-box-and-youll-still-pay-500

16. How It Works: Xbox Kinect. (n.d.). Retrieved October 1, 2019, from https://www.jameco.com/jameco/workshop/howitworks/xboxkinect.html

17. Modules - Kinetisense. (n.d.). Retrieved October 1, 2019, from Kinetisense website: https://www.kinetisense.com/modules/

18. Range of Motion Movement Catalogue. (n.d.). Kinetisense.

19. Sensortec, B. (2014). Intelligent 9-axis absolute orientation sensor. BNO055 Datasheet, November. Retrieved from https://www.robot-electronics.co.uk/files/BNO055.pdf

Study on a Method of Resolving Common Misjudgment by Color Sensors of Robots used in Education

Edison Suzuki
The American School in Japan; Tokyo, Japan

***Abstract* – LEGO® MINDSTORMS® EV3 is a robot kit that is often used for the introduction of Science, Technology, Engineering and Mathematics (STEM) education to students around the world. Using the color sensor and color detection mode that are standard on the EV3 and trying to determine the color of color blocks, the color discrimination was not successful every time due to individual differences of each color sensor. In the standard color identification mode of EV3, there were many misjudgments of color blocks, especially between red and yellow blocks. Therefore, another method for identifying a color block was considered in this study. RGB mode was used for color identification. G/B ratio was calculated for blue and green block identification, and R/G ratio was calculated for red and yellow block identification. These values did not overlap between the blue and green blocks and between the red and yellow blocks. The intermediate value of G/B ratio between blue and green blocks and intermediate value of the R/G ratio between red and yellow blocks were used to correctly distinguish the two blocks. As a result, the color blocks could be identified as almost 100%.**

Key Words - color misjudgment, color sensor, LEGO® MINDSTORMS® EV3, RGB mode.

INTRODUCTION

LEGO® MINDSTORMS® EV3 is a robot kit that is often used for the introduction of Science, Technology, Engineering and Mathematics (STEM) education to students around the world. The configuration of the robot includes one robot brain, three motors, and five sensors including the color sensor. However, using the color sensor and color detection mode that are standard on the EV3 and trying to determine the color of color blocks, the color discrimination was not successful every time due to individual differences of each color sensor.

The color identification by the EV3 uses the "color identification" mode that comes standard with the EV3 itself. The "color identification" mode are planned to identify colors into 7 colors [1][2]. However, even if robots with EV3 were made with the same configuration using the "color identification" mode as default settings, each EV3 could not identify color blocks perfectly.

At the time of the release of EV3, only this "color identification" mode was used for color detection. However, a new RGB (R: red, G: green, B: blue) mode program for EV3 has been released in recent years [3]. Therefore, the RGB mode program for advanced color identification was installed into the EV3 and was used to try to identify colors.

The RGB mode of the optical sensor detects the detected light with the three primary colors (R: red, G: green, B: blue) of the light, and identify the target color [4]. After measuring each R value, G value, and B value of the color block using the RGB mode, it was possible to identify each color block by a certain method. As a result, this RGB mode led to participation in the World Robot Olympiad 2018 Japan National Convention (WRO September 2018; Elementary School Expert Section) (Figures. 1 and 2). A method of resolving common misjudgment of each color block using the RGB mode is reported in this study.

FIGURE 1
SCENE OF WORLD ROBOT OLYMPIAD (WRO) 2018
JAPAN NATIONAL CONVENTION (KANAZAWA)

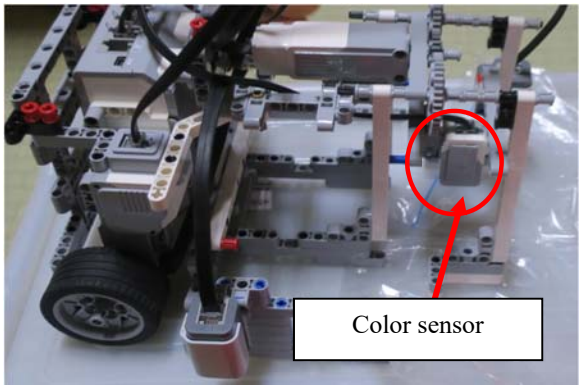


FIGURE 2
APPEARANCE OF THE EV3 ROBOT FOR WRO2018

TABLE 1
PATTERNS OF JUDGMENT BY COLOR SENSORS

	—	×	○	○
	×	—	○	○
	○	○	—	×
	○	○	×	—

LITERATURE REVIEW

- EV3 standard color identification mode:

The standard color sensor of EV3 uses Red/ Green / Blue light in order with full-color LEDs on the object and reads the amount of reflection with a photodiode. The color sensor output a value according to the color it detected (0: no color, 1: black, 2: blue, 3: green, 4: yellow, 5: red, 6: white, 7: brown) [1][2].

- Three primary colors of light and RGB color space:

The RGB is one of the color spaces that represent colors mathematically [4]. The RGB color space is mostly used for computer monitors. Red, green, and blue are the three primary colors of light [4].

- EV3 RGB color identification mode:

A function for obtaining RGB raw values with a color sensor for EV3 was released (2016) [3]. By using RGB values, patterns of color identification can be more in an environment where the actual EV3 is placed.

METHODOLOGY

I. Recognition of misjudgment patterns generated in the standard color identification mode of EV3

When the color detection mode of EV3 standard equipment was activated, the misjudgment of blue blocks and green blocks, and the misjudgment of red blocks and yellow blocks occurred (Table 1). The incidence of misjudgment was greater for red and yellow blocks. Therefore, there has been a demand for a method that can reliably perform color identification in two groups of blue and green blocks, and red and yellow blocks.

II. Download RGB mode software

The educational version of LEGO® MINDSTORMS® EV3 (LEGO System A / S, Denmark) was made to download software that can use the RGB mode [1].

RESULTS

I. Measurement of RGB values of each color block by RGB mode

Before running the EV3 body on the practice course, RGB values of each color block (blue, green, red and yellow) were measured twice on the same practice course. Then, the RGB values of each color block at that time were manually read and recorded (n=46).

II. Consideration of regularity for color identification using RGB values

1. Calculation of G value-to-B value (G/B) ratio of blue and green blocks

The average RGB values of each color block (blue, green, red, and yellow) were shown in Figure 3. Looking at the RGB values of the blue and green blocks, it was found that there was almost no difference between the R and G values. On the other hand, the B value was found to be about 1/4 of the green block compared to the blue block. Based on these facts, it was considered that B value was effective to distinguish blue and green blocks. However, setting the threshold using only B value did not distinguish blue and green blocks correctly.

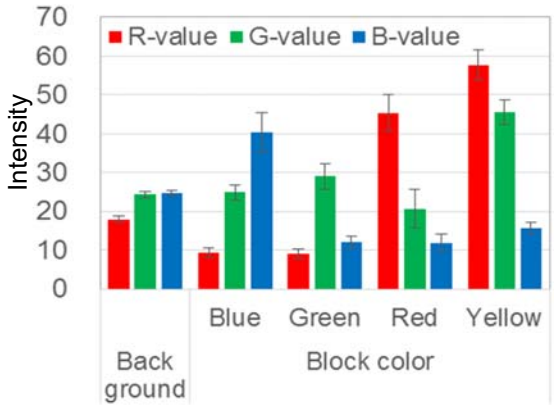


FIGURE 3
AVERAGE INTENSITY OF RGB VALUES OF EACH COLOR BLOCK

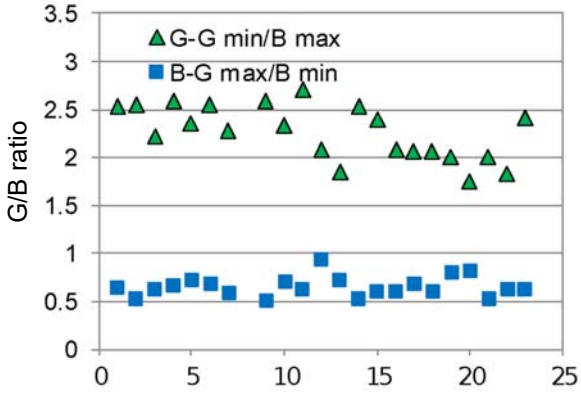


FIGURE 4
G/B VALUES OF BLUE AND GREEN BLOCK

Therefore, G value-to-B value (G/B ratio) as a threshold for the judgment of blue and green blocks was examined. When calculating the G/B ratio, the highest ratio in the blue block and the lowest ratio in the green block were considered as in (1) and (2). In (1), max (G) represents the highest value of G, and min (B) represents the lowest value of B. In (2), min (G) represents the lowest value of G, and max (B) represents the highest value of B. Each G/B ratio data of blue blocks and green blocks were plotted into Figure 4.

Blue block;

$$\text{G/B ratio} = \max(G)/\min(B) \quad (1)$$

Green block;

$$\text{G/B ratio} = \min(G)/\max(B) \quad (2)$$

From Figure 4, it was found that the G/B ratio was 0.94 or less for the blue block and 1.75 or more for the green block. There was a gap that showed the G/B ratio did not overlap between both blocks.

Therefore, to distinguish between blue blocks and green blocks, the threshold of the G/B ratio was set about 1.2. When the G/B ratio is <1.2, the block is judged as blue. When the G/B ratio is 1.2 <, the block is judged as a green block.

2. Calculation of R value-to-G value (R/G) ratio of red and yellow blocks

Looking at the RGB values of the red block and the yellow block (Figure. 3), the tendency of R value> G value> B value is common to both, and the B value is almost the same for both blocks. However, the red block tended to have a larger difference between the R and G values than the yellow block. From these facts, the R value-to-G value (R/G ratio) seemed to be effective in distinguishing red and yellow blocks.

When calculating the R/G ratio, the lowest ratio in the red block and the highest ratio in the yellow block were considered as in (3) and (4).

Red block;

$$\text{R/G ratio} = \min(R)/\max(G) \quad (3)$$

Yellow block;

$$\text{R/G ratio} = \max(R)/\min(G) \quad (4)$$

In (3), min (R) represents the lowest value of R, and max (G) represents the highest value of G. In (4), max (R) represents the highest value of R, and min (G) represents the lowest value of G. Each R/G ratio data of red blocks and yellow blocks were plotted into Figure 5.

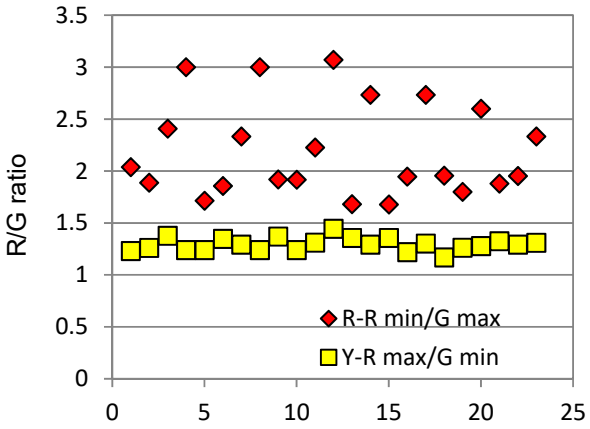


FIGURE 5
R/G VALUES OF RED AND YELLOW BLOCK

Figure 5 showed that the R/G ratio was 1.44 or less for the yellow block and 1.71 or more for the red block. It was found that the R/G ratio between these blocks does not overlap with each other. Therefore, to distinguish between a red block and a yellow block, the threshold of the R/G ratio was set about 1.5. When the R/G ratio is <1.5, the block is judged as yellow. When the R/G ratio is 1.5<, the block is judged as a red block.

As described above, as a result of calculating the G/B ratio and the R/G ratio and using them as the color identification threshold values of blue and green and red and yellow, the color blocks could be identified almost 100%.

DISCUSSION

The misjudging between the red block and the yellow block was more frequent than the misjudging between the blue block and the green block with EV3 standard color identification mode (Table 1). This phenomenon might have been occurred because of the wide distribution of the R/G ratio of the red block, from 1.68 to 3.07. Compared with the G/B ratio (Fig. 4), the R/G ratio of red and yellow blocks were sometimes very close (Fig. 5). Therefore, it is most difficult to distinguish red and yellow blocks.

In this study, it was found that the EV3 can be operated in RGB mode, and color identification can be made more accurately by setting the original threshold value for each color block. This means that it is important to understand the individuality and characteristics of each sensor. In addition to the original method for identification of colors, positions of installation of sensors of EV3 and timing of color identification are also important to identify each color. Thus it is very important to make an entire EV3 robot body from a view point of proper operation of each sensor.

Moreover, there was a tendency that the RGB values slightly differed even when the operating time with the same battery was short or the charging time of the battery was different. Therefore, it is important for participating in a robot tournament to determine the charging time and preparation for stable conditions on the robot every time. In the near future, the relationship between RGB values and the battery charging time will be studied.

CONCLUSION

In the standard color identification mode of EV3, there was a misjudgment of color blocks, especially between red and yellow blocks. Therefore, the RGB mode was used for color identification. After measurement of the RGB value of each block, the G/B ratio of blue and green blocks, and the R/G ratio of red and yellow blocks were calculated. The threshold of the G/B ratio and R/G ratio were considered from these data and used into the program. As a result, the color blocks could be identified as almost 100%.

ACKNOWLEDGMENTS

The author wishes to thank Dr. T. Kawasaki for her continuous support and guidance in conducting the experiments.

REFERENCES

[1] https://afrel.co.jp/archives/15674, “EV3 color sensor usage example (in Japanese).” (Sept. 15nd, 2019)

[2] https://afrel.co.jp/archives/15670, “Characteristics of EV3 color sensor (in Japanese).” (Sept. 15nd, 2019)

[3] https://afrel.co.jp/archives/20491, “Get RGB value with color sensor (in Japanese).” (March 2nd, 2019)

[4] Jonathan Sachs, Digital image basics; 1996, Digital Light & Color p.10

End-to-end Classification of Ballroom Dancing Music Using Machine Learning

Noémie Voss and Phong Nguyen
Sevenoaks School, Sevenoaks, United Kingdom
Email: voss02@sevenoaksschool.org, phong.nguyen@tokyotechies.com
* The two authors contributed equally to this work

Abstract **- 'Ballroom dancing' is a term used to designate a type of partnered dancing enjoyed both socially and competitively around the world. There are 10 different types of competitive ballroom dancing, each performed to different styles of music. However, there are currently no algorithms to help differentiate and classify pieces of music into their distinct dance types. This makes it difficult for beginner and amateur ballroom dancers to distinguish pieces of music, and know which type of dance corresponds to the music they are listening to. We proposed using an end-to-end machine learning approach to help classify music into different types with efficient and high accuracy. We evaluated four machine learning models and found that a Deep Neural Network with three hidden layers is the model with highest accuracy of 83%. As a result, ballroom dancers will have an easier method of distinguishing between specific types of ballroom dancing music.**

Key Words - ballroom dancing, classification, deep neural network, machine learning

INTRODUCTION

Ballroom dancing is a term used to designate a type of partnered dancing enjoyed both socially and competitively in dance festivals such as the Blackpool Dance Festival [1]. These dances are split into distinct categories. 10 of these are categorised as competitive ballroom dancing; these include Chacha, Foxtrot, Jive, Paso Doble, Quickstep, Rumba, Samba, Tango, Viennese Waltz and Waltz.

Each of these dances corresponds to a specific type of music. Despite their variations, it can be challenging for the human ear to distinguish their characteristics and classify pieces of music into a specific dance category.

However, this process of discrimination can be greatly facilitated using machine learning. A model can be created to aid professional and amateur dancers to easily match a piece of music to its corresponding dance type using knowledge about these unique characteristics. This can be of use when it is necessary to find a song for a specific event, or when curious about which type of dance to dance to a song.

Machine learning is a popular technology nowadays, utilised to help build applications which are able to address a variety of similar problems. Classification models can be trained using machine learning, and applied to new data. Using a data-driven approach, a machine can be made to learn the characteristics of various types of ballroom dance music, and hence be able to classify them. This way, we can avoid having to handcraft rule-based processing for each individual dataset and dance type; rather, the machine can learn to classify music accurately and efficiently.

In this paper, we propose a method to classify ballroom music using a machine learning approach and evaluate different models, including Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Random Forest, and Deep Neural Network (DNN). We wish to embed this method into a mobile application to help beginners learn ballroom dancing more easily. With a public dataset consisting of 3992 pieces of audio data, we found that the classification accuracy of the models we tested are 60%, 50%, 76% and 83% respectively.

RELATED WORK

In this section, we briefly describe similar, already existent, work using machine learning in music classification.

In recent years, machine learning has been applied widely for classification tasks, especially in the signal processing field. The most popular task is to classify music into different categories, one that has been studied since the early days of the internet. Tzanetakis and Cook (2002) addressed this problem with supervised machine learning approaches such as Gaussian Mixture model and k-Nearest-Neighbour classifiers [6]. Three sets of features utilised in this task were timbral structure, rhythmic content and pitch content. Support Vector Machines (SVM) were introduced later by Scaringella and Zoia (2005) [7]. With the recent success of Deep Neural Networks, a number of studies applying these techniques to speech and other forms of audio data were undertaken, such as by Gemmeke et al. (2017) [8]. However, representing audio in the time domain for input to neural networks is not very straight-forward, because of the high sampling rate of audio signals. Therefore, feature engineering for audio data is still a more popular technique for audio signals. H. Bahuleyan (2018) has extracted both frequency features and time-domain features, and used Deep Neural Network (DNN) to classify music genre with high accuracy [9].

U. Marchand, G. Peeters (2016) [10] have proposed a method to classify music into different ballroom dancing types by representing music data based on the application of Scale Transform along the two dimensions of time and frequency. However, this technique does not use a machine learning approach, but rule-based classification.

MUSIC AND SOUND DATA

Sound data consists of a wave converted into an electrical signal. Firstly, the sound wave is translated into an analog signal. An analog signal is a continuous representation of a sound wave and can consist of any values. After being converted into an analog signal, the signal is then converted into a digital signal using an analog-digital converter (ADC), allowing the sound to be represented in a way that can be stored digitally. Digital audio consists of a continuous sequence of numerical samples. A digital signal is formed by using the analog signal to capture digital values which represent the amplitude of the signal.

When working with sound waves, it is important to know that frequency is equal to the pitch of the sound. A higher frequency results in a higher pitch. Frequency is measured in Hz (hertz) units. The human hearing range is from 20-20kHz. To capture all frequencies that humans can hear in an audio signal, ADCs sample recordings at a frequency that is approximately double the human hearing range, at a rate of 44,100Hz. [2]

FEATURES ENGINEERING FOR SOUND DATA

Data Collection and Preprocessing

To create and train a model to develop an accurate method of matching a particular piece of music to a specific type of dance, an initial requirement is the collection of data.

We used a public dataset called the Extended Ballroom dataset [11] containing ballroom dancing music excerpts extracted from the website www.ballroomdancers.com. The dataset consists of 4180 tracks, corresponding to 13 different dance types.

TABLE 1
NUMBER OF TRACKS IN EXTENDED BALLROOM DATASET

Ballroom Dance Category	Number of Audio Tracks
Chacha	455
Jive	350
Samba	468
Viennese Waltz	252
Waltz	529
Tango	464
Rumba	470
Quickstep	497
Foxtrot	507
Paso Doble	53
Salsa	47
Slow Waltz	65
West Coast Swing	23
Total	**4180**

We decided to omit Salsa, Slow Waltz and West Coast Swing, because they are not performed competitively. We also omitted Paso Doble, as the dataset of 53 tracks not sufficient to train the model and would have caused an imbalance. The final dataset consisted of 3992 tracks.

In order to increase the number of music samples for the machine learning algorithms, the data was cut into smaller segments, called 'windows': 10 seconds per piece, with 5 seconds of overlap. This also allowed the size of each piece of music to be standardised, making it easier to extract statistical features.

Feature Extraction

A total of 37 features, listed below, were extracted from each track. The first feature was the BPM (beats per minute, or tempo), which indicates the speed of the music. Each dance type has a specific range of tempos at which they are danced.

TABLE 2
BPM OF DIFFERENT CATEGORIES OF BALLROOM DANCING [3]

Ballroom Dance Category	BMP range
Chacha	120-128
Jive	168-184
Samba	96-104
Viennese Waltz	174-180
Waltz	84-90
Tango	120-140
Quickstep	200-208
Rumba	100-108
Foxtrot	112-120
Paso Doble	120-124

The second feature extracted was the time signature (count). The time signature of a piece of music is equivalent to the number of beats per measure, indicating the rhythm of the music. The main significance of this feature is to facilitate the differentiation between Waltz/Viennese Waltz, which has a time signature of 3/4, and the other categories of dance music, which have a time signature of either 2/4 or 4/4.

A third feature extracted from the pieces of music was the fingerprint record of the song. A fingerprint consists of a list of five values combined to form one hashtag. These values are the most common frequencies of the piece of music from within different frequency ranges. In a sound bite, frequency is the determinant of pitch. Therefore, the fingerprint of a song can be interpreted as a series of pitches, which appear most frequently within their frequency ranges, within the song. The five ranges were: 30-40Hz, 41-80Hz, 81-120Hz, 121-180Hz, and 181-300Hz. [4]

The final 34 features extracted further increased the ability of the final model's classification. Both short-term and mid-term features were extracted from each piece of music. Short term features can be described as features extracted from the piece of music after having split it into several short-term windows. Mid-term features can be described as certain statistics extracted from the short-term features, such as the mean and the standard deviation.

TABLE 3
FEATURES EXTRACTED

	Name of feature	Description of feature
1	Zero Crossing Rate	The rate of sign-changes of the signal during the duration of a particular frame.
2	Energy	The sum of squares of the signal values, normalized by the respective frame length.
3	Entropy of Energy	The entropy of sub-frames' normalized energies. It can be interpreted as a measure of abrupt changes.
4	Spectral Centroid	The center of gravity of the spectrum.
5	Spectral Spread	The second central moment of the spectrum.
6	Spectral Entropy	Entropy of the normalized spectral energies for a set of sub-frames.
7	Spectral Flux	The squared difference between the normalized magnitudes of the spectra of the two successive frames.
8	Spectral Rolloff	The frequency below which 90% of the magnitude distribution of the spectrum is concentrated.
9-21	MFCCs	Mel Frequency Cepstral Coefficients form a cepstral representation where the frequency bands are not linear but distributed according to the mel-scale.
22-33	Chroma Vector	A 12-element representation of the spectral energy where the bins represent the 12 equal-tempered pitch classes of western-type music (semitone spacing).

Classification Models

Waltz and Viennese Waltz have a different time signature to the other types of ballroom music, and therefore our classification model only uses time signatures as the main criteria to classify Waltz and Viennese Waltz from the other categories of dance. A second, separate, model is used to classify to which of the other seven types the song belongs to: Tango, Samba, Chacha, Rumba, Quickstep, Foxtrot Blues, or Jive.

The overall process (illustrated in Figure 1) results in a program able to classify music into the nine categories of ballroom dancing music.

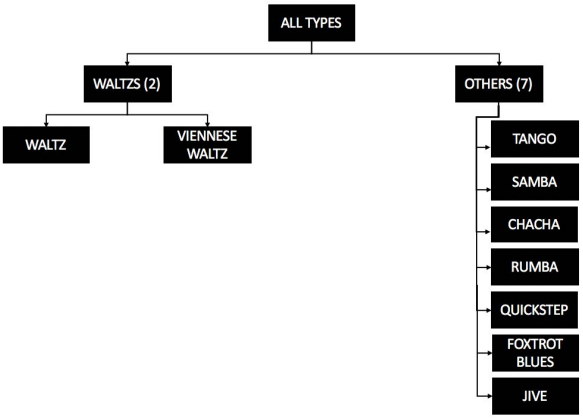


FIGURE 1
LAYOUT OF THE CLASSIFICATION

To classify the other seven types of ballroom dance music and Waltz or Viennese Waltz, we compared seven types of classification models: Support Vector Machines (SVM), k-Nearest Neighbors, Random Forest and Deep Neural Network.

Support Vector Machines (SVM) is a machine learning algorithm used for classification. The SVM model looks at the training data and creates an optimal decision boundary, known as a 'hyperplane', between the extremes of the different classes of data. The hyperplane separates the data into the different categories. [12]

A k-Nearest Neighbors model classifies data based on a number (k) of its nearest neighboring data points, regardless of their label. The unidentified data point is categorised based on the label of the majority of its neighbors, i.e., it takes the same label as the label which appears the most. [13]

A Random Forest model uses decision tree algorithms to create several decision trees which can classify data into discrete categories. Each separate tree predicts a label for the unidentified data point. The final label is the label with the most predictions. [14]

A Deep Neural Network consists of machine learning algorithms similar to the human brain. These algorithms are able to recognise patterns based on the features of the data. Using hidden layers, the model is able to classify each of the data points. [15]

Figure 2 below describes how we processed data to train a machine learning model to classify different ballroom dance types and evaluate the model.

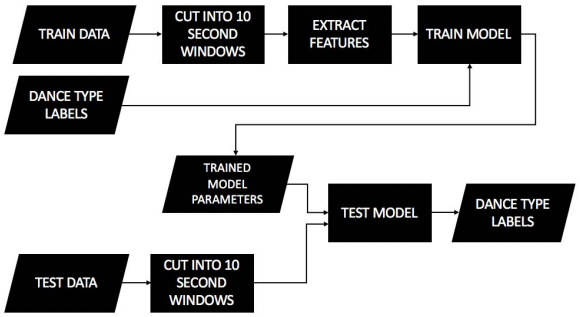


FIGURE 2
OVERVIEW OF THE PROCESS

Classification of a New Song

To use our method to classify a new song, we preprocessed the audio tracks by cutting them into 10-second windows with a five second overlap. For each standardized 10 second piece, after all features are extracted, we utilised time signature feature to group the pieces with a time signature of three as the first group (for Waltz and Viennese Waltz), and the second group of all other seven types of dance.

In the final output, each 10 second window piece belonging to a song had a classification label. We determined the ballroom dance label of the song based on the majority characterization.

EXPERIMENTAL SETTINGS

The first step taken in creating the model was collecting the necessary data. We collected the data from the public Extended Ballroom dataset. As mentioned above, we removed the tracks for four types of dance: Paso Doble, Slow Waltz, Salsa and West Coast Swing. The created a final data set of 3992 tracks.

In order to evaluate our method, we split the data randomly with 10% of the data as the test data, and 90% of the data as the train data. In addition, we also tested the model on our own collected dataset with songs chosen by professional dancers. We picked 10 songs for each type of dance.

After the preprocessing, the total number of 10-second pieces was 45000 for the train data and 4100 for the test data. We then extracted the aforementioned 37 features from each piece and built four different classifiers: SVM, k-NN, Random Forest, and DNN. We evaluated each model in terms of accuracy and weighted F1-score. Hyperparameters of each model were tuned to their best performance by trial-error processes.

The models were not only evaluated on the test data, but also on a dataset of 30 professional ballroom dancing songs.

RESULTS

Several different types of classification models were trained and evaluated in order to maximise the accuracy of the final model chosen. These models were: 'Support Vector Machine', 'K-Nearest-Neighbours' with neighbours=101, 'Random Forest' with 100 trees, and a 'Deep Neural Network' with three hidden layers.

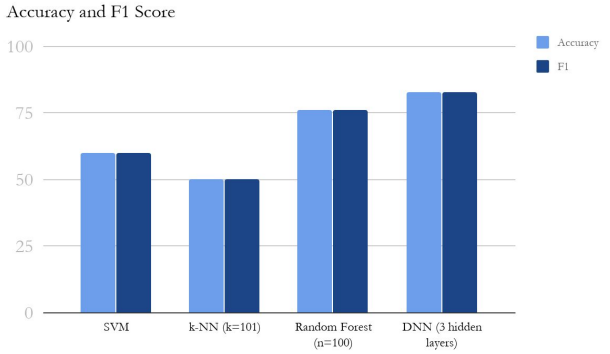


FIGURE 3
A GRAPH COMPARING THE F1 SCORES AND ACCURACIES OF EACH TYPE OF CLASSIFICATION MODEL

The accuracy and F1 scores of the models are specified in Table 4.

TABLE 4
ACCURACY AND F1 SCORE OF EACH MODEL

Model	Accuracy	F1 Score
SVM	60%	60%
k-NN (n=101)	50%	50%
Random Forest (n=100)	76%	76%
DNN (3 hidden layers)	80%	80%

The results showed that the DNN model with 3 hidden layers has the highest accuracy. When we tested it on a dataset of 30 songs for ballroom dancing, it could recognize 27 songs accurately, which translated to 90% accuracy.

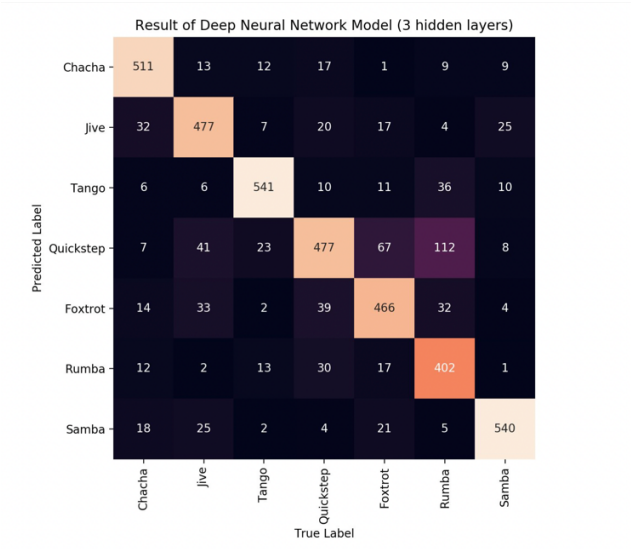


FIGURE 4
CONFUSION MATRIX OF DEEP NEURAL NETWORK MODEL WITH THREE HIDDEN LAYERS FOR THE 'EXTENDED BALLROOM' DATASET

Figure 4 above is a confusion matrix to show the results of the Deep Neural Network model with three hidden layers when tested on the Extended Ballroom dataset of 3992 tracks. The model is accurate at classifying both Tango and Chacha. However, the model is not able to classify Rumba and Quickstep as well as the other types of dance; 112 Rumba tracks were incorrectly labelled as Quickstep.

The overall accuracy of the model is 83%.

Once programmed into a mobile application, beginners in ballroom dancing will be able to use this application to improve their ability of distinguishing between various music, hence improving their confidence in ballroom dancing. This application can also be gamified to encourage beginners to guess the dance category before checking their answers with the model, providing a fun way for dancers to improve their ability to distinguish between dance types. The actual mobile application has not been implemented yet, but could be feasible in the future.

CONCLUSION

We have evaluated an end-to-end method using machine learning to classify a piece of music belonging to a specific type of ballroom dancing. The Deep Neural Network with three hidden layers showed an accuracy of 83%. We open our source-code of the processing to train the model at: https://bitbucket.org/xphongvn/ballroom_music_classification/

One way to improve this method will be finding more relevant sound features that can help classify ballroom dancing music with higher accuracy. An alternative way is to have more training data. We plan to publish a bigger dataset for ballroom dancing music in the future

REFERENCES

1. Blackpool Dance Festival, Website, Accessed in 2018 August, <http://http://www.blackpooldancefestival.com/>
2. Jovanovic, J, "How does Shazam work? Music Recognition Algorithms, Fingerprinting, and Processing", Website, Accessed in 2018 September, <http://https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-r ecognition>
3. Beats Per Minute Online, "Tempo Indications and Beats Per Minute (BPM) Reference for Social Dance Genres", Website, Accessed in 2018 June, <http://http://www.beatsperminuteonline.com/en/home/bpm-beats-per-minute-reference-for- dance-genres>
4. Jovanovic, J, "Music Recognition: Fingerprinting a Song", Website, Accessed in 2018 Sep, <http://http://https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting- and-recognition>
5. Giannakopoulos, T, "3. Feature Extraction", Website, Accessed in 2018 August, <https://github.com/tyiannak/pyAudioAnalysis/wiki/3.-Feature-Extraction>
6. George Tzanetakis and Perry Cook. 2002. Musical genre classification of audio signals. IEEE Transactions on speech and audio processing 10(5):293– 302.
7. Nicolas Scaringella and Giorgio Zoia. 2005. On the modeling of time information for automatic genre recognition systems in audio signals. In ISMIR. pages 666–671.
8. Jort F Gemmeke, Daniel PW Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R Channing Moore, Manoj Plakal, and Marvin Ritter. 2017. Audio set: An ontology and human-labeled dataset for audio events. In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, pages 776–780.
9. Bahuleyan, H. (2018). Music genre classification using machine learning techniques. arXiv preprint arXiv:1804.01149.
10. Marchand, U., & Peeters, G. (2016, September). Scale and shift invariant time/frequency representation using auditory statistics: Application to rhythm description. In 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1-6). IEEE.
11. Marchand, U., & Peeters, G. (2016). The extended ballroom dataset.
12. Ray, Sunil, and Business Analytics. "Understanding Support Vector Machine Algorithm from Examples (along with Code)." Analytics Vidhya, 11 Mar. 2019. www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/
13. Srivastava, Tavish. "Introduction to KNN, K-Nearest Neighbors : Simplified." Analytics Vidhya, 7 Mar. 2019, www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/ .
14. Koehrsen, Will. "Random Forest Simple Explanation." Medium, Medium, 27 Dec. 2017, medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d.
15. "A Beginner's Guide to Neural Networks and Deep Learning." Skymind, skymind.ai/wiki/neural-network.

Classical Cryptography Methods: Security Improvements by the Hill Cipher

Rainata Vania Hemawani
Ritsumeikan Uji, Kyoto, Japan
Email: rainatavania@gmail.com

Abstract - **The Hill Cipher is a classical cryptography, as there is no present usage of its original method. Nevertheless, it contributed in advancing the mathematics in cryptography at the time it was invented. It was the first method that allowed encryption of more than 2 letters at once, as it utilizes matrix to transform plaintexts into cipher texts. In this paper, the improvements of Hill Cipher will be explored by comparing it with an older method, the Caesar Cipher, looking at how the Hill Cipher managed to reduce the flaws in its method. Furthermore, the advantage of the use of matrices in the cryptosystem of Hill Cipher is going to be explored.**

Key Words – Hill Cipher, Caesar Cipher, Classical Cryptography, Matrix.

INTRODUCTION

1.1 Background

Cryptography is the science of creating and decoding secret messages [1]. It has been practiced since around 1900 BC when the first evidence was found in the form of an inscription carved in Egypt [2]. The diagram below shows how it works.

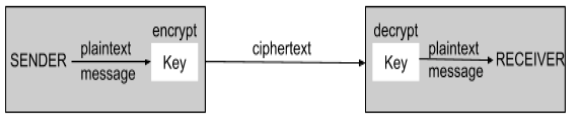


FIGURE 1
A DIAGRAM OF A CRYPTOSYSTEM [2]

A sender wants to send the receiver a message [3]. The original form of the message is the plaintext [4]. To secure the message, the plaintext must be turned into a cipher text [3], the transformed message, through a process called encryption [4]. A key is used in this process [3]. The key is the most important aspect in a cryptosystem [3]. Only the sender and receiver should have access to it. The cipher text will be sent to the receiver, but the cipher text itself is usually accessible to public, thus it has to be secured [3]. Decryption is a process done by the receiver using the key to convert the cipher text back to its plaintext from to obtain the message [4].

Classical cryptography methods are no longer used in the present society [3]. This method was done without the use of computers. Third parties who did not have access to the key yet were willing to decode the message went through a process called cryptanalysis (discussed in a latter section). This process led to the development of mathematics as methodological flaw in the methods are found through code-breaking.

This paper will show the security improvements made by a method called the Hill Cipher by looking at its advantages at that time, in comparison with an older method. The Hill Cipher was a polygraphic substitution cipher, which was a method that reduced the possibility of manually code-breaking [5]. The encryption, decryption and cryptanalysis before the polygraphic substitution cipher were heavily more reliant on language abilities to crack the code [5]. The Hill Cipher, as a part of polygraphic substitution cipher contributed in making cryptography transition into mathematics [5].

How did the Hill Cipher, a polygraphic substitution cipher, introduced in 1929 improve upon the existing monoalphabetic cipher methods?

THE HILL CIPHER

Hill Cipher is a polygraphic substitution method, allowing encryptions of a large set of alphabets [6]. Concepts like matric multiplications and matrix inverses are applied in this method [6]. To introduce the encryption method, the sentence below is going to be used as plaintext:

WAS THIS SENTENCE EASY TO CRACK

The plaintext will be converted into column matrices [7], where alphabets will be converted into numbers, like such, seen below:

A	**B**	**C**	**D**	**E**	**F**	**G**
0	1	2	3	4	5	6
H	**I**	**J**	**K**	**L**	**M**	**N**
7	8	9	10	11	12	13
O	**P**	**Q**	**R**	**S**	**T**	**U**
14	15	16	17	18	19	20
V	**W**	**X**	**Y**	**Z**		
21	22	23	24	25		

Normally, the whole plaintext will be encrypted all at once, making the dimension of our plaintext 26×1, since the sentence has 26 letters. However, to simplify the process of encryption and decryption, our given plaintext will be encrypted in pairs (WA, ST, HI, etc), where each pair will have the dimension of 2×1. These 2×1 matrices will be represented as the variable P in any encryption or decryption explanations. Here are all pairs of plaintexts converted into matrices:

$\binom{W}{A} = \binom{22}{0}$ $\binom{T}{E} = \binom{19}{4}$ $\binom{0}{C} = \binom{14}{2}$
$\binom{S}{T} = \binom{18}{19}$ $\binom{N}{C} = \binom{13}{2}$ $\binom{R}{A} = \binom{17}{0}$
$\binom{H}{I} = \binom{7}{8}$ $\binom{E}{E} = \binom{4}{4}$ $\binom{C}{K} = \binom{2}{10}$
$\binom{S}{S} = \binom{18}{18}$ $\binom{A}{S} = \binom{0}{18}$
$\binom{E}{N} = \binom{4}{13}$ $\binom{Y}{T} = \binom{24}{19}$

Since the plaintext matrices are 2×1, the key matrix (K) will be a square matrix of 2×2 [7]. A key matrix also has to fulfil the requirement of $\gcd(determinant, 26) = 1$ [8]. The elements in the key matrix will be integers within the range $0 - 25$, as the calculations are working in $mod26$. The key matrix that will be used is:

$$K = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix} \quad [9]$$

The cipher text, presented as variable C, are the letters that are going to be sent to the receiver. The formula below is used to encrypt each plaintext:

$$C = KP mod 26 \quad [10]$$

Matrix K (the key) will be multiplied with each pair of plaintext, P.

$$\begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 22 \\ 0 \end{pmatrix} = \begin{pmatrix} 66 \\ 44 \end{pmatrix}$$

All values should be presented in its mod26 form:

$$\begin{pmatrix} 66 \\ 44 \end{pmatrix} mod26 = \begin{pmatrix} 14 \\ 18 \end{pmatrix}$$

The rest of the plaintexts are encrypted into its cipher text below:

$$C_2 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 18 \\ 19 \end{pmatrix} mod26 = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$$
$$C_3 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 7 \\ 8 \end{pmatrix} mod26 = \begin{pmatrix} 19 \\ 2 \end{pmatrix}$$
$$C_4 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 18 \\ 18 \end{pmatrix} mod26 = \begin{pmatrix} 4 \\ 22 \end{pmatrix}$$
$$C_5 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 4 \\ 13 \end{pmatrix} mod26 = \begin{pmatrix} 25 \\ 21 \end{pmatrix}$$
$$C_6 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 19 \\ 4 \end{pmatrix} mod26 = \begin{pmatrix} 17 \\ 6 \end{pmatrix}$$
$$C_7 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 13 \\ 2 \end{pmatrix} mod26 = \begin{pmatrix} 19 \\ 10 \end{pmatrix}$$
$$C_8 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 4 \\ 4 \end{pmatrix} mod26 = \begin{pmatrix} 24 \\ 2 \end{pmatrix}$$
$$C_9 = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ 18 \end{pmatrix} mod26 = \begin{pmatrix} 2 \\ 12 \end{pmatrix}$$
$$C_{10} = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 24 \\ 19 \end{pmatrix} mod26 = \begin{pmatrix} 25 \\ 13 \end{pmatrix}$$
$$C_{11} = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 14 \\ 2 \end{pmatrix} mod26 = \begin{pmatrix} 22 \\ 12 \end{pmatrix}$$
$$C_{12} = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 17 \\ 0 \end{pmatrix} mod26 = \begin{pmatrix} 25 \\ 8 \end{pmatrix}$$
$$C_{13} = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 2 \\ 10 \end{pmatrix} mod26 = \begin{pmatrix} 10 \\ 2 \end{pmatrix}$$

The matrix form of the cipher texts is listed below:

$$\begin{pmatrix} 14 \\ 18 \end{pmatrix}\begin{pmatrix} 7 \\ 1 \end{pmatrix}\begin{pmatrix} 19 \\ 2 \end{pmatrix}\begin{pmatrix} 4 \\ 22 \end{pmatrix}\begin{pmatrix} 25 \\ 21 \end{pmatrix}\begin{pmatrix} 17 \\ 6 \end{pmatrix}\begin{pmatrix} 19 \\ 10 \end{pmatrix}$$
$$\begin{pmatrix} 24 \\ 2 \end{pmatrix}\begin{pmatrix} 2 \\ 12 \end{pmatrix}\begin{pmatrix} 25 \\ 13 \end{pmatrix}\begin{pmatrix} 22 \\ 12 \end{pmatrix}\begin{pmatrix} 25 \\ 8 \end{pmatrix}\begin{pmatrix} 10 \\ 2 \end{pmatrix}$$

The matrices above are converted back into alphabets, in the same form of our initial plaintext:

OSHBTCEWZVRGTKYCCMEMWMZIKC

EARLIER METHODS

To understand how the Hill Cipher was an improvement at the time it was invented, we will compare it with a simple substitution cipher, like the Caesar Cipher.

3.1. Caesar Cipher

Caesar cipher is a monoalphabetic cipher only allowing encryption of one alphabetic character at a time [10]. It only requires basic concepts such as addition and subtraction. The same plaintext used in the Hill Cipher encryption is going to be used as an example of Caesar cipher encryption.

WAS THIS SENTENCE EASY TO CRACK

Since Caesar Cipher only involved addition and subtraction, the key will be a chosen number where $K \in \mathbb{Z}^+$, and $1 \leq K \leq 25$, as $mod26$ is going to be applied. The encryption formula to find the cipher text for the Caesar cipher is:

$$C = (P + K)\ mod26 \quad [11]$$

The key in Caesar Cipher determines how many letters away would the cipher text be from the plain text (based on alphabetical order). By choosing a number randomly from 1 to 25, the key that is going to be used is:

$$K = 5$$

The way the alphabets are converted into numbers is the same as the conversion method of Hill Cipher. The plaintext in alphabetical from is converted into numbers below:

$P =$ 22 0 18 19 7 8 18 18 4 13 19 4 13 2 4
4 0 18 24 19 14 2 17 0 2 10

The formula will be applied to each letter of the plaintext because it is a monoalphabetic method. The working process to get the cipher text is presented:

$(22+5)\ mod26$ $(19+5) mod26$ $(18+5)\ mod26$
$(0+5)\ mod26$ $(7+5)\ mod26$ $(18+5)\ mod26$
$(8+5)\ mod26$ $(8+5)\ mod26$ $(13+5)\ mod26$
$(18+5)\ mod26$ $(19+5)\ mod26$
$(4+5)\ mod26$
$(13+5)\ mod26$
$(2+5)\ mod26$
$(4+5)\ mod26$

$(4+5)\ mod26$ $(19+5)\ mod26$ $(2+5)\ mod26$
$(0+5) mod26$ $(14+5) mod26$ $(17+5)\ mod26$
$(18+5)\ mod26$ $(0+5)\ mod26$
$(24+5)\ mod26$ $(2+5)\ mod26$
$(10+5)\ mod26$

$$\therefore C = 1\ 5\ 23 \quad 24\ 12\ 13\ 23 \quad 23\ 9\ 18\ 24\ 9\ 18\ 7\ 9$$
$$9\ 5\ 23\ 3 \quad 24\ 19 \quad 7\ 22\ 5\ 7\ 15$$

The numerical values above are then converted back into alphabets, resulting in the cipher text below:

BFXYMNXXJSYJSHJJFXDYTHWFHP

4.1. Cryptanalysis

The process of breaking codes to leak information, and finding weaknesses of a cryptographic method is called *cryptanalysis* [12]. The Caesar Cipher is vulnerable to "cipher text only attack", which is a cryptanalysis method where the plaintext can be found by only having access to the cipher text [13].

4.1.1 Frequency Analysis

Frequency analysis refers to decoding a cipher text, knowing which alphabets appear more frequently [14]. This depends on the language of which the plaintext was written in [14]. Since it is a monoalphabetic substitution, it implies that a letter that appears often in the cipher text, will be decrypted as a letter that appears as often in the plaintext. The plaintext was originally written in English. The most frequently appearing letters in English include the letter "E" and "T" [14]. We can assume that the frequently appearing letters in the cipher text are likely to be one of the most frequently appearing letters in English. Table 1 shows the frequency of the letters in the cipher text:

TABLE I
FREQUENCY OF ALPHABETS IN THE CIPHER TEXT

B	F	X	Y	M	N	S	J	H	D	T	W	P
1	3	3	3	1	1	3	4	2	1	1	1	1

The letter "J" in the cipher text appears most often. We will assume that it is the vowel "E". So, we assume that $E \rightarrow J$ after encryption. We will substitute what we know into the encryption formula of Caesar Cipher, as shown below:

$$J = (E + K)\ mod26$$

Making "J" and "E" into their numerical forms in the equation will give us:

$$9 = (4 + K)\ mod\ 26$$

With $9-4$, we will get 5. Since $mod26$ is applied before getting our cipher text to keep the range within 0-25, there are infinite possibilities for the value of K.

$$K = 5, 31, 57 \ldots$$

However, the use of $mod26$ will make the value back to 0 once it reaches 26, so using any of the values above will lead us to the same value for P. Thus, we can assume that it is 5 instead of the other bigger values. In this case, we got the correct answer for the key. By figuring out the right key, the whole cipher text can be broken.

4.1.2 Brute Force

The cryptanalysis method previously heavily relies on alphabetical frequency [13], but letter frequency is not always the case. However, due to the simplicity of Caesar Cipher, "cipher text only" cryptanalysis can be done by listing all possibilities. The key needed to decrypt is only in the form of a single number and there are only 25 possibilities for the key.

We have calculated that $WAS \rightarrow BFX$ when $K = 5$. However, since there were no spaces given in the cipher text, a decoder will not know that the plaintext starts with a three-letter word, thus let us assume that the decoder will try breaking the first four letters, $BFXY$. The first letter of the cipher text is "B". Listing all the possibilities will allow also allow us to list all the possible keys, as shown in Table 2:

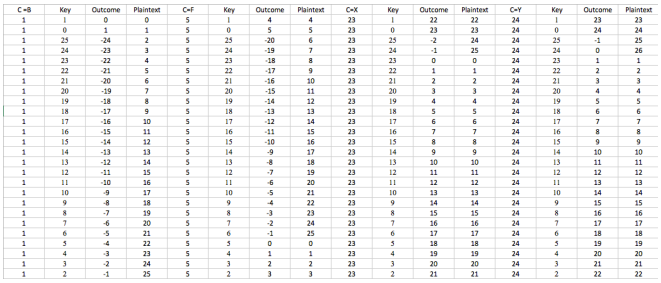
TABLE 2
POSSIBLE PLAINTEXTS AND KEYS FOR "B"

Possible Plaintexts	Key		
C	25	Q	11
D	24	R	10
E	23	S	9
F	22	T	8
G	21	U	7
H	20	V	6
I	19	W	5
J	18	X	4
K	17	Y	3
L	16	Z	2
M	15	A	1
N	14		
O	13		
P	12		

Finding the plaintext is done very similarly, follow the formula below:

$$P = (C - K)\ mod 26 \qquad [11]$$

In Figure 2 below, cipher text $BFXY$ was subtracted with all the possible keys. This was done using Microsoft Excel for convenience, but is also very possible to be done manually:



C=B	Key	Outcome	Plaintext	C=F	Key	Outcome	Plaintext	C=X	Key	Outcome	Plaintext	C=Y	Key	Outcome	Plaintext
1	1	0	0	5	1	4	4	23	1	22	22	24	1	23	23
1	0	1	1	5	0	5	5	23	0	23	23	24	0	24	24
1	25	-24	2	5	25	-20	6	23	25	-2	24	24	25	-1	25
1	24	-23	3	5	24	-19	7	23	24	-1	25	24	24	0	26
1	23	-22	4	5	23	-18	8	23	23	0	0	24	23	1	1
1	22	-21	5	5	22	-17	9	23	22	1	1	24	22	2	2
1	21	-20	6	5	21	-16	10	23	21	2	2	24	21	3	3
1	20	-19	7	5	20	-15	11	23	20	3	3	24	20	4	4
1	19	-18	8	5	19	-14	12	23	19	4	4	24	19	5	5
1	18	-17	9	5	18	-13	13	23	18	5	5	24	18	6	6
1	17	-16	10	5	17	-12	14	23	17	6	6	24	17	7	7
1	16	-15	11	5	16	-11	15	23	16	7	7	24	16	8	8
1	15	-14	12	5	15	-10	16	23	15	8	8	24	15	9	9
1	14	-13	13	5	14	-9	17	23	14	9	9	24	14	10	10
1	13	-12	14	5	13	-8	18	23	13	10	10	24	13	11	11
1	12	-11	15	5	12	-7	19	23	12	11	11	24	12	12	12
1	11	-10	16	5	11	-6	20	23	11	12	12	24	11	13	13
1	10	-9	17	5	10	-5	21	23	10	13	13	24	10	14	14
1	9	-8	18	5	9	-4	22	23	9	14	14	24	9	15	15
1	8	-7	19	5	8	-3	23	23	8	15	15	24	8	16	16
1	7	-6	20	5	7	-2	24	23	7	16	16	24	7	17	17
1	6	-5	21	5	6	-1	25	23	6	17	17	24	6	18	18
1	5	-4	22	5	5	0	0	23	5	18	18	24	5	19	19
1	4	-3	23	5	4	1	1	23	4	19	19	24	4	20	20
1	3	-2	24	5	3	2	2	23	3	20	20	24	3	21	21
1	2	-1	25	5	2	3	3	23	2	21	21	24	2	22	22

FIGURE 2
APPLYING BRUTE FORCE USING EXCEL

The outcome is the result of subtracting the cipher text with the key. Since some values become negative, $C = Outcome + 26$ because $mod26$ has to be applied to fulfil $0 \leq C \leq 25$. The cipher text in Figure 2 is in number form, so it is converted into alphabets, as shown in the table below:

TABLE 3
POSSIBLE PLAINTEXTS

Key	Cipher Text	Key	Cipher Text
25	CGYZ	12	PTLM
24	DHZA	11	QUMN
23	EIAB	10	RVNO
22	FJBC	9	SWOP
21	GKCD	8	TXPQ
20	HLDE	7	UYQR
19	IMEF	6	VZRS
18	JNFG	5	WAST
17	KOGH	4	XBTU
16	LPHI	3	YCUV
15	MQIJ	2	ZDVW
14	NRJK	1	AEWX
13	OSKL		

If a decoder assumed that the plaintext was in English, the most plausible cipher text is "WAST", highlighted in Table 3, as they could guess that it might have started with the word "WAS". Thus, the decoder can deduce that the key is most likely 5. This allows the codebreaker to decode the rest of the cipher tex. This was a weakness of Caesar Cipher, as one of the past methods before the Hill Cipher came about.

SECURITY

The Hill Cipher managed to reduce, if not eliminate, the possibility of frequency analysis. Let us compare the plaintext with both encrypted texts:

Plaintext:	WAS	THIS	SENTENCE	EASY	TO	CRACK
Caesar Cipher:	BFX	YMNX	XJSYJSHJ	JFXD	YT	HWFHP
Hill Cipher:	OSH	BTCE	WZVRGTKY	CCME	MW	MZIKC

In the Caesar Cipher, all plaintext "E" will definitely turn into J if $K = 5$. Moreover, all "S"s turn into "X"s. This feature makes it very possible to do frequency analysis, because by knowing the plaintext for a single alphabet allows the decoder to crack the whole cipher text. However using the Hill Cipher, the "E"s in the plaintext become a variety of letters like "Z", "G", "Y" and "C". It is true that the cipher text of Hill Cipher was encrypted by putting them in pairs, but even if a plaintext is encrypted all at once, a plaintext that appears frequently would most likely have multiple cipher texts. Taking one of the pair from the plaintext as an example:

$$P = \begin{pmatrix} E \\ E \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

There was one pair that repeats the letter "E". The encryption process of this pair will show that the cipher text does not become a repeated alphabet:

$$P = \begin{pmatrix} 4 \\ 4 \end{pmatrix}, K = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix} mod 26 = \begin{pmatrix} 24 \\ 2 \end{pmatrix}$$

$$\therefore EE \rightarrow YC$$

The encryption of Hill Cipher uses matrix multiplication [5]. It is an advantage and an improvement from monoalphabetic ciphers. Since matrix multiplication is used for encryption, the decryption involves finding the inverse of a matrix. The use of matrices made brute force way more difficult to do. Let us look through the decryption process of the Hill Cipher. Here below is the formula for decryption:

$$P = \{K^{-1}\} C mod 26$$

$\{K^{-1}\}$ refers to the inverse of the key matrix, but in order to find P, all the values are should be in their $mod26$ form, including K^{-1}. Thus, the process below shows how fractions and negative values are replaced into positive integers. We know that:

$$C = OS = \begin{pmatrix} 14 \\ 18 \end{pmatrix}$$

$$K = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}$$

A key matrix is written in the form of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $determinant = ad - bc$. Thus, for our chosen key matrix, the determinant will be:

$$3(5) - 3(2) = 9$$

Originally, we would find an inverse of a 2×2 matrix by using the given formula:

$$K^{-1} = \frac{1}{det}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

The result needs to be converted back into alphabets, therefore the elements in matrix P should be within the range of $0 - 25$. Multiplying the adjoint $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ with the constant $\frac{1}{det}$ will make the elements in matrix P possibly fractions. Thus, the constant becomes the mod26 form of $\frac{1}{det}$. It is known that $\frac{a}{b}(modn) \equiv k$ such that $bk \equiv a(modn)$ [15]. k will be the constant we will multiply the adjoint matrix with. Therefore:

$$\frac{1}{9}mod26 \equiv k \mid 9k \equiv 1mod26$$
$$27 \equiv 1mod26$$
$$9(3) \equiv 1mod26$$
$$\therefore k = 3$$

Next, the elements in the adjoint should also be in their mod26 form to change the negative values into positive. The adjoint of $\begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}$ is $\begin{pmatrix} 5 & -3 \\ -2 & 3 \end{pmatrix}$, however, the adjoint that is going to be used to find $\{K^{-1}\}$ will be $\begin{pmatrix} 5 & -3mod26 \\ -2mod26 & 3 \end{pmatrix}$ which results to $\begin{pmatrix} 5 & 23 \\ 24 & 3 \end{pmatrix}$.

We have found both the constant and adjoint to find $\{K^{-1}\}$. Using the formula of finding the inverse of a 2×2 matrix, we will then find $\{K^{-1}\}$, while making all values into their $mod26$ form:

$$3\begin{pmatrix} 5 & 23 \\ 24 & 3 \end{pmatrix} = \begin{pmatrix} 15 & 69 \\ 72 & 9 \end{pmatrix}$$

$$\{K^{-1}\} = \begin{pmatrix} 15 & 69 \\ 72 & 9 \end{pmatrix} mod26 = \begin{pmatrix} 15 & 17 \\ 20 & 9 \end{pmatrix}$$

In order to find the plaintext, we will substitute the matrices $\{K^{-1}\}$ and C into the decryption formula:

$$\begin{pmatrix} 15 & 17 \\ 20 & 9 \end{pmatrix}\begin{pmatrix} 14 \\ 18 \end{pmatrix} mod26 = \begin{pmatrix} 22 \\ 0 \end{pmatrix} = WA$$

We can confirm that it is correct as the first two letters of out plaintext was indeed WA.

From the example above, we have found that $\{K^{-1}\} = \begin{pmatrix} 15 & 17 \\ 20 & 9 \end{pmatrix}$. $\{K^{-1}\}$ has the same characteristics as a key matrix, thus, $\{K^{-1}\}$ can also be written as matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where the elements are within the range of $0 - 25$.

Let us look at a case where a cryptanalysis of the Hill Cipher is attempted to be done, where the ciphertext is known, but the key matrix is unknown. Substituting the known ciphertext to the formula $P = \{K^{-1}\}Cmod26$, we will get the following, where matrix $\begin{pmatrix} x \\ y \end{pmatrix}$ represents the plaintext:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 14 \\ 18 \end{pmatrix} mod26 = \begin{pmatrix} x \\ y \end{pmatrix}$$

We will get the following equations:

$$(14a + 18b)mod26 = x$$
$$(14c + 18d)mod26 = y$$

There are 6 variables to solve from the two equations, unlike in Caesar Cipher, where there were only 1 variable to solve since the key is in the form of a single value. Another similar approach is trying to solve by the encryption formula, but this does not eliminate the need of solving for 6 different variables. Let us substitute the ciphertext into the encryption formula $C = KPmod26$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} mod26 = \begin{pmatrix} 14 \\ 18 \end{pmatrix}$$

We will get the following equations:

$$(ax + by)mod26 = 14$$
$$(cx + dy)mod26 = 18$$

Let us represent $ax + by$ as Q and $cx + dy$ as R in this case. In order to solve this, first we would need to find the value of Q and R in the equation $Q\ mod26 = 14$ and $R\ mod26 = 18$ to find what $ax + by$ and $cx + dy$ could be. This is possible, as $mod26$ gives a limit to the possible value of x and y. However, the next step would be solving for 6 different variables.

Hill Cipher does not eliminate the fact that there are only 25 other possibilities of what a letter in the plaintext could be, in cases where the plaintext was written in roman alphabets. This means that the elements in the key matrix, including $\{K^{-1}\}$ will be in the range $0 \leq K \leq 25$, where K in that statement represents all the variable a, b, c, d in the key matrix. Nevertheless, the smallest possible key matrix would have a dimension of 2×2. In a 2×2 key matrix, there are 4 elements in it. Therefore, even though there are only 26 possible alphabets, there are 26^4 possible key matrices. This makes Hill Cipher secured from the possibility of manual brute force. If the plaintext "WAS THIS SENTENCE EASY TO CRACK" encrypted all at once

instead of in pairs, the key matrix will have a dimension of 26×26 so it makes 26^{26} possible key matrices. The encryption method of the Hill Cipher makes brute force more difficult, and the difficulty increases and the dimension of the key matrix increases.

CONCLUSION

Cryptography methods constantly improve over time, and cryptanalysis also develops as new, unbroken methods are brought about. The Hill Cipher was one of the methods that contributed in the advancement of cryptography mathematics, where the cryptanalysis of it no longer depends on language. As Hill Cipher was able to encrypt more than one letter at a time, it reduced the possibility of breaking the code using frequency analysis, especially when the dimension of the key matrix increases. In a Hill Cipher operation, matrices are used to group plaintexts and cipher texts, thus the key is in a form of a matrix. This characteristic of the Hill Cipher makes it very difficult to conduct cryptanalysis through brute force by hand. However, as technology and cryptography method improve over time, Hill cipher is now categorised as classical cryptography.

ACKNOWLEDGMENT

The author wishes to thank Mr. Gurnick for the guidance and feedbacks that contributed in the completion of this research.

REFERENCES

[1] Brown, Alyson. "The Basics of Cryptography." *Towards Data Science*, Towards Data Science, 28 Jan. 2019, towardsdatascience.com/the-basics-of-cryptography-80c7906ba2f7.

[2]Sidhpurwala, Huzaifa. "A Brief History of Cryptography." *Red Hat Customer Portal*, 14 Aug. 2013, access.redhat.com/blogs/766093/posts/1976023.

[3] "The Hill Cipher, A Linear Algebra Perspective." doi:https://facultyfp.salisbury.edu/despickler/personal/Resources/LinearAlgebra/HillCipherHandoutLA.pdf.

[4]"Introduction to Cryptography." *Classical Cryptography*, 1 Sept. 2010, people.eecs.ku.edu/~hossein/Teaching/Fa10/710/Readings/Classical-Cryptography.pdf.

[5]"Polygraphic Substitution Ciphers." *Crypto*, www.crypto-it.net/eng/simple/polygraphic-substitution-ciphers.html?tab=0.

[6]"Hill Cipher." *Crypto Corner*, crypto.interactive-maths.com/hill-cipher.html.

[7]Christesen, Chris. "Polygraphic Substitution Cipher: The Hill Cipher." NKU, 2015, www.nku.edu/~christensen/1402%20Hill%20cipher.pdf.

[8]"Hill Cipher." *Practical Cryptography*, practicalcryptography.com/ciphers/hill-cipher/.

[9]"Hill Cipher Encryption 2by2 Matrix." *YouTube*, YouTube, 2 Aug. 2017, www.youtube.com/watch?v=GY-psGvlKgA.

[10]Holden, Joshua. *The Mathematics of Secrets: Cryptography from Caesar Ciphers to Digital Encryption*. Princeton University Press, 2019.

[11]Panigrahy, Saroj Kumar, et al. "Image Encryption Using Self-Invertible Key Matrix of Hill Cipher Algorithm." 21 Feb. 2007.

[12]"Caesar Cipher." *Practical Cryptography*, practicalcryptography.com/ciphers/classical-era/caesar/.

[13]Edgar, Thomas W, and David O Mans. "Cryptanalysis." *ScienceDirect* , 2017, www.sciencedirect.com/topics/computer-science/cryptanalysis.

[14]"Frequency Analysis: Breaking the Code." *Crypto Corner*, crypto.interactive-maths.com/frequency-analysis-breaking-the-code.html.

[15]Eisenberg, Murray. "Hill Ciphers and Modular Linear Algebra." 3 Nov. 1999.

Text Classification Based on Machine Learning: Quantifying the Sentiment of Online Reviews

Shreya Subramaniam
The American School in Japan
Tokyo, Japan
20subramaniams@asij.ac.jp

***Abstract*—Machine learning technologies and AI make performing crucial tasks much more efficient, which betters people's lives. Examples of this can be seen in every aspect of our lives in our day-to-day interactions with technology such as facial recognition in smartphones and complex alarm systems to keep houses safe. All such programs improve the efficiency and accuracy of important tasks through data collection. Nowadays, machine learning and AI has proved to be incredibly useful when gathering large amounts of data that would otherwise require immense amounts of human effort. It can be applied in a variety of ways, but one especially lucrative application of such technology is in the field of advertising and market research. Large companies require information about their consumers and target demographic to better their products and services.This study aims to use sentiment analysis techniques, such as Naive Bayes, to determine the way customers feel about corporations.**

***Index Terms*—machine learning, data science, sentiment analysis, text classification, AI**

I. Introduction

Artificial Intelligence (AI) technologies are rapidly becoming prevalent in how we do business, manage money, gather data, and function in society. With their quickly advancing capabilities, AI and Machine Learning technologies are becoming more and more useful for our daily life. Automated systems on smartphones, security cameras, and home appliances all utilize AI and machine learning technologies to assist us in making our lives safer and easier by completing tasks in a faster or smarter fashion, or by completing tasks that humans are incapable of doing themselves. In the research process, it is particularly useful for analysis of immense quantities of information. One such example of this is using gathered data to determine the sentiment and positivity of reviews written by users of products or services.

When creating and running a business, understanding and learning from consumer feedback is imperative to establishing a loyal consumer base as well as attracting new customers. Basically humans can read and classify reviews and feedbacks of customers in text format, but the number of reviews or feedback a person can classify in a certain amount of time is limited in either quantity or accuracy. If we have some solution that can learn and classify the reviews and feedback automatically, we can overcome this limitation by making the classification process automatic, thus classifying reviews faster. With the recent advancements of AI and Machine Learning technology, and with numerous available sources of data, such as social media and review websites, it is possible to have some solution for this problem. This motivates us to research on how to employ AI and Machine Learning technologies for this problem. In this paper, we focus on investigating data sources and Machine Learning theory/algorithms to build up Machine Learning models that classify the degree of positivity in reviews and feedback in text format.

Through extensive research, we gathered reviews from popular websites, such as Yelp and Tripadvisor, built up Machine Learning models and used the data to train the models in order to quantify their level of positivity. This automates and facilitates the analyzing of customers' criticisms and concerns. Organizations or companies will be able to understand the feedback and reviews better to better satisfy customers' demands, in order to increase profits.

In order to conduct our research, we developed and trained a sentiment analysis system to parse through and tokenize reviews data from review websites such as Tripadvisor and Yelp. Sentiment analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral [1]. In essence, the system extracts specific words from each review and quantifies them based on the degree of satisfaction it reflects from the consumer's point of view. It then collects the quantities assigned to each of the reviews and creates an average – one that is the result of an accumulation of data from several different websites, all into one program.

Ultimately, the increased efficiency provided by data collection and analysis leads to an increase in knowledge on both the consumer and company's part. Consumers will get a more accurate representation of the quality of a company's services or products. Corporations will be able to more accurately cater to their audience and make necessary reforms in shorter periods of time, increasing revenue.

II. Background

In total, we used three different types of classification methods to analyze the sentiments present in online hotel reviews. These three programs are Naive Bayes, Support Vector Machine (SVM), and K Nearest Neighbor (KNN). Each classification method, while similar in nature and outcome, operates on a set of distinct principles in order to effectively

identify positive and negative sentiments, as well as the degree of positivity or negativity.

A. Naive Bayes Classifier

This model was built and trained using the Naive Bayes classifier, a popular model for the classification and categorization of text since the 1960's. It assumes that all data points are independent of one another in order to conduct an analysis of their relationship and ultimately determine the location and nature of a desired point of information. The Naive Bayes classification system operates chiefly on the Bayes Theorem:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

The Naive Bayes classifier (also known as Simple Bayes or Independence Bayes) assumes that the probability of event A, given that B has already occurred, is the same as the probability of event A regardless of whether or not B has occurred. This strong independence assumption between events is what gives the Naive Bayes classifier the "naive" name. Naive Bayes is one of simple probabilistic classifiers. These classifiers analyze a range of data and create a distribution based on the probability that a data point could fall into a range of classes, rather than merely spitting out a single class that the data point most likely belongs to. These types of classifiers have become integral parts of the text classification process in many ways due to their accuracy and usefulness in their own right, as well as combined with other classifiers.

First introduced to the text retrieval community in the early 1960's, the Naive Bayes classifier has remained a popular baseline method for text categorization because it uses qualities such as word frequency to judge whether or not a text belongs to a certain category (ie. spam vs inbox, positive vs negative, etc.). This has many applications in the real world, such as making automatic medical diagnoses and doing sentiment analysis for companies.

B. Support Vector Machine

Similarly to KNN, Support Vector Machine analyzes the way words in a review are grouped together to create a certain level of positivity or negativity in a review rather than quantifying sentiment based on individual words. When used in conjunction with methods like KNN and Naive Bayes, SVM allows us to analyze both the sentiment expressed in one review as well as the total positivity or negativity of sentiments expressed in all reviews pertaining to a certain company.

Specifically, SVM operates in a series of planes. It groups together similar data in one plane and creates an ideal hyper plane that houses the predictions the program makes based on the input. It categorizes several groups of data as distinct entities and predicts what a new data point will look like by comparing its characteristics to those of previously existing points or planar groups.

C. K Nearest Neighbor

Often referred to as KNN, K Nearest Neighbor analyzes each review in relation to the others similar to it, whether that be in date written, sentiment expressed, etc. This text classification method allows us to look at reviews in relation to others to gain a more holistic view of the ratio of positive and negative sentiment to total reviews.

KNN, unlike SVM, is a non-parametric approach to data classification. Both SVM and KNN cluster data based on the points similar to it. However, KNN makes predictions about new data points by comparing the new data point to previously existing points as individual entities, rather than the groups they belong to. In this respect, it employs similar independence assumptions that Naive Bayes does while still taking into account the ways in which common behavioral trends impact the accuracy of predictions.

III. Approach

The main method we used to tokenize the sentiments of people's opinions surrounding large corporations was sentiment analysis. We used methods such as Naive Bayes to tokenize and quantify the level of positivity present within reviews. More specifically, we obtained data from Yelp, Twitter, and Tripadvisor to train and test a sentiment analysis model that would assign numbers to comments.

These numbers are from a -1 to 1 scale, with 0 representing neutral reviews. Essentially, what this program did was take review samples in the dataset that were given a certain rating and used that to calibrate itself. By feeding a large amount of data into the program, we were able to quantify the sentiments expressed in the reviews in a more holistic, standardized way that represented the sentiments of reviews across several reviewing platforms.

A. Datasets

We used mass amounts of data from several different review websites to train and test a model that can accurately quantify the sentiments of reviews. In the datasets we used, we primarily focused our attention on the relationship between the "review text" column and "star rating" column. In total, our dataset contained over 10,000 reviews. To collect data, we searched different review websites and social media platforms for different review types. On websites such as Tripadvisor and Yelp, reviewers are able to write reviews, as well as put a numerical rating on a scale from 1 to 5 (1 being the worst level of customer satisfaction and 5 being the best) for the reviews. However, many social media platforms do not operate in the same way. Due to this discrepancy in the types of data we had to collect, we decided to train our model using a variety of data types, including positive and negative sentiment outputs, so that the model could be able to recognize positive and negative sentiment present in a multitude of different data presentations. This way, we can make sure that our data is representative of all the opinions regarding the performance of a certain product or company. To combat the difference in review style, we took

into account factors such as hashtags and mentions, as those are fairly common features across all social media websites.

We collected all our data manually to gather the same amount of data from a diverse set of sources, which increases our ability to conclude the effectiveness of our representation.In addition to this, we also gathered data from various free databases that contain free-for-use, public datasets. These datasets were created by others for the data science community to use in their research endeavors, as cited below.

The dataset we used to train and test our model was the Hotel Reviews set from Datafiniti at the Data World website. In total, it contains 25 unique features and 10,000 observations. These features include the contents of the review, the review title, the number of stars assigned to each review, and the name of the hotel. When we were searching for potential datasets, we wanted to make sure we chose one that would allow us to train our model on reviews from multiple different hotels and hotel chains so that it could be as holistic as possible. In addition, we recognize that factors such as the title of the review could be an indicator of sentiment in the same way that stars and review text conveys sentiment. In addition to this, the dataset also contains important demographic information which could potentially allow us to determine outside factors that may influence how customers feel about certain companies. Although we started out using this dataset to train and test our model, we will be combining this data with other sets to train our model to analyze sentiments for a large variety of corporations.

To build and train the model, we wrote several different programs, each using a different type of classifier, to calculate the sentiment of the hotel reviews. In all of our programs, we split the dataset into "train" and "test" categories in order to familiarize our program with the type of data we needed to use. We did this so we could analyze the different results given by each type of classifier and determine whether there were any inconsistencies and mistakes with the data itself or with the results of the other code. In addition, using multiple types of classifiers allowed us to analyze the different output formats and helped us determine which classifier would be most useful for us and the most easy to read and work with.

B. Method

Our program aimed to tokenize the words in online reviews to associate them with a particular star value. To do this, we used several different classification methods, each one having similar structures and functions to the others. Firstly, we loaded a preexisting dataset and split it into a train and test group. The train group, composed of approximately 75 percent of the data, was used to help the program associate common words and phrases with a certain level of positivity or negativity. By doing this, our program was able to more accurately and precisely predict the star rating of a certain review based on the words it has. The test group of the dataset was comprised of 25 percent of the data. This was meant to see how well the program could identify and tokenize certain words as well as how close to the actual star rating the program was. After we used one dataset to train and test all three models, we

IV. Evaluation

In order to evaluate our results, we cross-checked our programs with each other, making sure that they all used the same datasets and that they all used different classification methods. In total, we were able to create programs from SVM, Naive Bayes, and KNN. All three of those classification methods revolve around different presentations of results, therefore allowing us to examine our results in different ways and seeing which output formats could be ideal for presenting our findings to others, especially those who would greatly benefit from our program. When we cross-examined our data, we expected the results to say the same things. Generally, if they do, this means that there are no major errors with our program or our methods of analysis. However, this step is incredibly useful because errors that come up in cross-examination could indicate inconsistencies with the data we used or flaws in our programs.

A. Classification Report

In order to test the sensitivity, accuracy, and precision of our models, we created a classification report for each program that would give us specific insight into the effectiveness of our predictions. For all programs, we were able to get an f1 score of over 80 percent (0.8) for all the datasets we used. This tells us that our programs were accurate and precise across the board. In addition, it tells us that we can use our programs to accurately and precisely predict the rating given to a particular review.

TABLE I
Naive Bayes

Class	F1 Score	Precision	Recall
Review	0.85	0.87	1.00
Stars	0.84	0.94	1.00

This is the classification report printed from the program using Naive Bayes to classify the reviews. Overall, the program was able to precisely identify key words and phrases in reviews 87 percent of the time, and was subsequently able to use that information to consistently predict the star rating 94 percent of the time. These numbers are relatively higher than what would occur in a reliable program (about 60 to 65 percent). The Naive Bayes program was also able to recall the different positive and negative key words nearly 100 percent of the time. Generally, it is impossible to say that a program is 100 percent effective. However, we used about 10,000 reviews to train and test our program, which is a relatively small sample size compared to all the reviews on the Internet. Due to this, we were able to confidently say that all the program was able to correctly recall and implement all the positive and negative phrases we programmed it to understand. The program's F1 score is an indicator of accuracy and is calculated by the following equation:

$$F_1 = (\frac{recall^{-1} + precision^{-1}}{2})^{-1}$$

Essentially, an F_1 score of 85 percent for review accuracy and 84 percent for star prediction accuracy shows that our program is very accurate and precise when determining positive and negative sentiment of reviews and subsequently, star predictions.

TABLE II
SUPPORT VECTOR MACHINE (SVM)

Class	F1 Score	Precision	Recall
Review	0.81	0.96	1.00
Stars	0.87	0.91	1.00

The SVM classification method proved to be more effective in both F_1 score and precision than Naive Bayes. This could be a result of the fact that Naive Bayes assumes that all data points are independent, while SVM doesn't make that same assumption.

TABLE III
K NEAREST NEIGHBOR (KNN)

Class	F1 Score	Precision	Recall
Review	0.89	0.91	1.00
Stars	0.88	0.92	1.00

The KNN classification method produced results with an effectiveness comparable to both the SVM method and Naive Bayes method. Due to the fact that all three programs gave results that were similar down to the hundreth digit, there is virtually no statistical significance to the differences in hundreths of percentage points between the three programs.

B. Root Mean Square Error (RMSE)

Given that our data are categorical, we cannot use Root Mean Square Error (RMSE) or R-squared methods to cross-examine or assess the validity of our findings. Both of those methods revolve around calculating the residuals of the data, which we cannot do due to the fact that each observation we study is qualitative, even though the final sentiment ranking output is a number. Ultimately, the number's significance comes from the conversion of initial qualitative inputs to a value that represents the abstract notion of satisfaction and contentment. While the number may not outwardly display the different factors that contributed to it, it is an important part of summarizing a business' overall performance and can potentially lead to businesses' growing more self-aware and more willing to make internal changes that will further customer satisfaction.

V. CONCLUSION

This paper sought to explore the applications of sentiment analysis and text classification in the context of corporate profit by training and testing multiple algorithms to quantify the positivity of customer hotel reviews across a variety of review websites. To do so, we included multiple data sets and multiple types of classification methods to ensure that our results would be precise and accurate. In total, we analyzed tens of thousands of hotel reviews from platforms such as TripAdvisor and Yelp from hotels all over the world.

By exploring the different methods of text classification, we were able to determine the relative effectiveness of certain classification methods at predicting and analyzing the relationship between key words and positivity levels.

Through learning about and using many different methods of sentiment analysis, we were able to harness the benefits of all of them to analyze the different ways positive and negative sentiment can express itself in an online review, thus helping us determine how positively people felt about certain hotel companies and chains.

ACKNOWLEDGMENTS

The author would like to thank Dr. Thien X. Phan for his aid as a research supervisor and Aneesh Chand for helping edit the paper.

REFERENCES

[1] Frank, E. and R. Bouckaert, R. (n.d.). Naive Bayes for Text Classification with Unbalanced Classes. [ebook] Computer Science Department, University of Waikato, New Zealand. Available at: https://www.cs.waikato.ac.nz/ eibe/pubs/FrankAndBouckaertPKDD06new.pdf [Accessed 22 Sep. 2019].

[2] "Scikit-learn.org. (2019). 1.9. Naive Bayes — scikit-learn 0.21.3 documentation." [online] Available at: https://scikit-learn.org/stable/modules/naivebayes.html" [Accessed 22 Sep. 2019].

[3] "1.9. Naive Bayes — scikit-learn 0.21.3 documentation", Scikit-learn.org, 2019. [Online]. Available: https://scikit-learn.org/stable/modules/naivebayes.html. [Accessed: 22- Sep- 2019].

[4] "Naive Bayesian", Saedsayad.com, 2019. [Online]. Available: https://www.saedsayad.com/naivebayesian.htm. [Accessed: 22- Sep- 2019].

[5] "How Naive Bayes Algorithm Works? (with example and full code) — ML+", Machine Learning Plus, 2019. [Online]. Available: https://www.machinelearningplus.com/predictive-modeling/how-naive-bayes-algorithm-works-with-example-and-full-code/. [Accessed: 22- Sep- 2019].

[6] "Naive Bayes Classifiers - GeeksforGeeks", GeeksforGeeks, 2019. [Online]. Available: https://www.geeksforgeeks.org/naive-bayes-classifiers/. [Accessed: 22- Sep- 2019].

[7] Manuscript Templates for Conference Proceedings", Ieee.org, 2019. [Online]. Available: https://www.ieee.org/conferences/publishing/templates.html. [Accessed: 22- Sep- 2019].

[8] "Sentiment Analysis using SVM", Medium, 2019. [Online]. Available: https://medium.com/@vasista/sentiment-analysis-using-svm-338d418e3ff1. [Accessed: 22- Sep- 2019].

[9] "Text Classification: A Comprehensive Guide to Classifying Text with Machine Learning — MonkeyLearn", MonkeyLearn, 2019. [Online]. Available: https://monkeylearn.com/text-classification/. [Accessed: 22- Sep- 2019].

[10] Lee, H. Nguyen, K. Lykowski and T. Phan, A Machine Learning Based Approach for Automated Used-Car Price Evaluation. International Young Researchers' Conference, 2018.

[11] A. Tripathy, A. Agarwal and S. Kumar Rath, Classification of Sentimental Reviews Using Machine Learning Techniques. Science Direct.

[12] W. Yang Wang, E. Lin and J. Kominek, This Text Has the Scent of Starbucks: A Laplacian Structured Sparsity Model for Computational Branding Analytics. Seattle, Washington, USA: Association for Computational Linguistics, 2013.

[13] I. Maks and P. Vossen, Sentiment Analysis of Reviews: Should we analyze writer intentions or reader perceptions?. Hissar, Bulgaria, 2013.

How can Expedia Group Inc. effectively use HomeAway brand to maintain its position as the market leader with Airbnb Inc. rapidly expanding their Online Travel Agency business?

Airi Ueda
Ritsumeikan Uji, Uji city, Kyoto
Email: airi.ueda7@gmail.com

***Abstract* – With the rapid growth of Airbnb, the current poison of Expedia in the OTA industry has been threatened because HomeAway has its strengths and weaknesses compared to Airbnb. Airbnb had continued to expand its target market by increasing its property listings. HomeAway's Unique selling point of lending the whole properties has no longer been the unique selling point, putting them into a competitive disadvantage. Regarding these changes that Airbnb has been making, Expedia needs to plan some strategies to encourage both hosts and guests to use HomeAway to compete with Airbnb.**

Keywords – Target Market, Brand Recognition, Unique Selling Point, Promotion, Marketing Mix

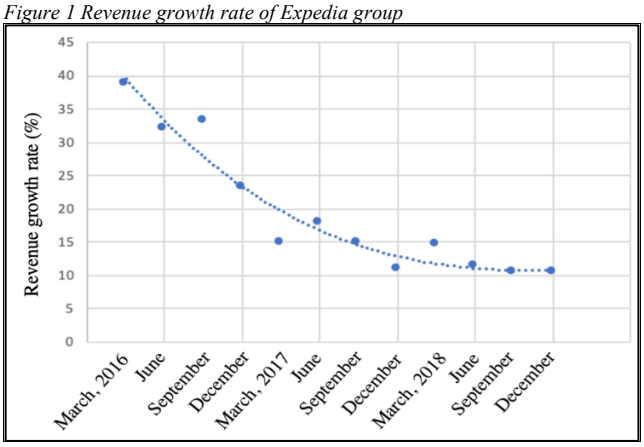
INTRODUCTION

Online travel agency (OTA), is an online travel website that specializes in the sales of a variety of travel products such as hotels, flights, and activities. Expedia Group is one of the world's leading online travel agency launched by Microsoft Corporation in 1996.[1] Expedia Group has expanded internationally and the brands they own covers a wide range of services. Many of the online travel companies also offer several types of services other than hotel bookings such as car rentals. Expedia's competitor, Booking Holdings currently owns more market share than Expedia. In fact, in 2017, Booking Holdings' total revenue was $12.68 billion while Expedia's total revenue was $11.22 billion.[2] However, when only focusing on the online hotel booking market, Expedia is the market leader since Expedia owns 19.52% of the online hotel booking market while Booking.com owns 13.29% of the market in 2017.[3]

Despite Expedia Group's success in becoming one of the leading online travel brands, recently with Airbnb rapidly growing and expanding their business, Expedia Group's position in the OTA industry has been threatened. A few years after Airbnb was founded, Expedia started to become more aware of the fast-growing new alternative accommodations industry, which is the vacation rental industry. Thus, they decided to enter the vacation rental industry by acquiring HomeAway in 2015.[4]

However, with Airbnb rapidly growing and Expedia failing to invest time and money until recently to improve the services HomeAway provided, HomeAway started to lose its market share.[3] Finally, during the last year, Expedia has started to put more effort into the HomeAway brand and invested more to raise their number of rental properties they offer.[5] As a result, HomeAway now generates about $300 million in revenue a quarter, which is about 10% of Expedia Group's total revenue.[6] Nevertheless, Airbnb is still the market leader in the vacation rental industry with revenue about $2.6 billion and HomeAway is still behind Airbnb.[6]

Figure 1 Revenue growth rate of Expedia group



Although Expedia Group is still expected to increase its revenue in the future, their revenue growth rate is slowing down. This is especially because of the decline in the growth of Expedia's home rental service, which includes the HomeAway brand.[6]

Regardless of the efforts, the Expedia group has put in the HomeAway brand, they have failed to slow the growth rate of Airbnb. According to Morgan Stanley Research, a survey taken towards more than 4,000 consumers of Airbnb from 4 countries including the US and UK in 2016, shows that 47 percent of respondents have replaced their hotel stay with Airbnb with 90% of them being satisfied with their stay.[7] Regarding their successfulness, in 2017, Airbnb was predicted to increase its revenue by $6 billion in 2020.[7]On the other hand, according to David Binder Research

conducted in 2018, only 23% of the respondents were satisfied with the customer service Expedia Group serves.[3] allowing Expedia to become vulnerable to the increased competition in the OTA industry. With the rapid growth of Airbnb preventing Expedia Group to further grow and expand their business, Expedia group's revenue growth has been declining the last 3 years as shown in Figure 1. Therefore, in this paper, the online travel market specifically the vacation rental industry will be explored, especially focusing on the effective measures Expedia group can take using the HomeAway brand to counter Airbnb's rapid growth. In order to do so, the marketing mix of HomeAway will be compared with Airbnb.

PROPERTIES (PRODUCTS)

HomeAway is older than Airbnb and was first established in 2006 while Airbnb was first found in 2008. Although they both provide similar services and have listings of properties worldwide in more than 190 countries, since HomeAway only provides rental for entire properties while Airbnb offers rental for several types of properties, there is a clear difference in the number of properties they offer. HomeAway currently offers around 2 million properties to the users while Airbnb offers more than 4 million properties.[5]

Table 1 Comparison of approximate number of properties in some cities in 2016[8]

City	HomeAway	Airbnb
Kissimmee	12,163	5,950
Rome	10,473	25,386
Paris	9,804	65,973
Davenport, FL	7,808	3,113
Barcelona	5,805	21,737
Panama City beach	5,289	2,091
London	5,248	49,533
Myrtle Beach	5,034	1,127
NewYork	4,237	12,043
Gulf Shores, AL	4,228	457
Los Angeles	3,578	8,367
Miami	3,510	6,941
Hilton Head Island	3,439	760
Breckenridge	2,931	1,631
Destin	2,792	545

Although Table 1 only shows a few cities out of many cities, when compared to Airbnb, HomeAway seems to have more listings that are resort-like properties. For example, in cities like Rome and Paris, the number of properties listed on Airbnb is significantly greater than HomeAway while in cities like Kissimmee and Davenport, HomeAway owns more properties than Airbnb. In fact, HomeAway's former CEO, Jon Gray, once explained that they focused on resort-like areas to increase sales because "most travelers think of resort destinations for the holidays".[9] This supports Table 1, where HomeAway's properties are mainly focused on resort areas such as Kissimmee, which is close to Disney World.

Moreover, in 2014, HomeAway has become the largest "Luxury-Focused Vacation Rental Website in the World".[5] From this, since renting luxurious entire properties are expensive and tend to be rented by larger groups of people, it can also be argued that HomeAway is targeting mostly the middle-upper class families. Nevertheless, compared to Airbnb, unlike Airbnb with a wide variety of properties and prices, with HomeAway's listings being focused on luxurious resort-like entire properties, they may have a limited customer base. Therefore, it can be said that, compared to Airbnb, HomeAway is likely to face difficulties expanding its market.

In addition, the following tables are the data collected from MarketMinder, which is a website that compares the rental properties of Airbnb and HomeAway in cities from all around the world. Nine cities were compared in total: six cities were chosen evenly from the higher rank cities to lower rank cities within the top 50 international tourist destination cities within the top 50 international tourist destination cities,[9] and three cities located in resort areas chosen from Table 1 which HomeAway had larger number of properties than Airbnb. This clarifies how Airbnb owns larger market size in many touristy areas around the world.

Table 2 Market size comparison in some of the cities in 2018[10]

City	Market size (%)		
	Airbnb	HomeAway	Both
London	82	12	6
Paris	95	4	1
Tokyo	95	5	0
New York	91	6	3
Sydney	86	7	7
Los Angeles	89	6	5
Davenport	43	47	10
Kissimmee	48	42	10
Gulf Shores	18	55	27

Table 3 Ratio of the type of property available for rent in some cities in 2018[11]

City	Ratio of the type of property available for rent(%)		
	Entire home	Private room	Shared room
London	61	38	1
Paris	86	13	1
Tokyo	68	24	8
New York	55	42	3
Sydney	65	33	2
Los Angeles	69	25	6
Davenport	99	1	0
Kissimmee	97	2	1
Gulf Shores	100	0	0

From Table 2, by comparing the number of properties from Airbnb and HomeAway located in six cities which are London, Paris, Tokyo, New York, Sydney, and Los Angeles, every city had more than 85% of the market owned by the Airbnb. This emphasizes how the number of properties on the housing lists greatly differs between Airbnb and HomeAway.

In addition, since HomeAway only rents out the entire houses, it may be more inconvenient for the users who are just wanting to rent a single small space. In fact, looking at Table 3, cities where Airbnb is owning most of the market tend to show a higher percentage of private rooms and shared rooms available in that market whereas, places like resort cities in the US where HomeAway and Airbnb shared almost

the same amount of properties, more than 97% of the rental properties available was an entire home. From this, it is obvious that Airbnb has a wider variety in the properties they rent. This shows how Airbnb and HomeAway have different targeted customers. Airbnb can target a wider range of customers with its wide price range as well as great variety in the types of properties they offer, from private rooms to entire properties. However, HomeAway targets middle-upper families who are looking for luxurious trips. This difference in targeted customers is also evident in their vision statement. While Airbnb aims to provide "Airbnb for everyone"[12], HomeAway aims "To give people the space they need to drop the distractions of everyday life and simply be together".[9]

Table 4 Average rental size in some cities measured in terms of number of rooms[11]

City	Average rental size (number of bedrooms)
London	1.7 bedrooms
Paris	1.1 bedrooms
Tokyo	1.4 bedrooms
New York	1.4 bedrooms
Sydney	1.9 bedrooms
Los Angeles	1.5 bedrooms
Davenport	5.2 bedrooms
Kissimmee	4.9 bedrooms
Gulf Shores	2.6 bedrooms

Moreover, the six touristy cities and the three resort cities in the US had different average rental size measured by the number of bedrooms. While the six cities had an average of 1.5 rooms, three resort cities had an average of 4.2 rooms. Although the six touristy cities also provided rentals of entire properties, the average rental size is significantly smaller than the three resort cities. This indicates that there was less demand for properties with a larger number of rooms in the 6 touristy cities suggesting how trips with smaller groups may be more common in touristy cities.

The difference in average rental size may also relate to the difference in the average length of stay. A data suggests that Airbnb users stay on an average of 6.4 nights in New York, 4.6 nights in London and 8.4 nights in Paris.[13] When comparing with HomeAway, with HomeAway having a wider variety of properties recommended for stays more than a month[5] may suggest that HomeAway property listings are not suited in areas that are highly ranked in the international tourist destination cities with higher demand of properties for a shorter stay.

Providing rentals for the entire house has been HomeAway's unique selling point until a few years ago. In fact, in 2016, they used the phrase "Don't share" as their marketing strategy to promote how HomeAway users can rest assured that the space they rented is completely theirs during the visit.[2] However, with Airbnb offering rentals of the whole house now, this has no longer become a unique selling point for HomeAway. This allows Airbnb with various property types available to have competitive advantage than HomeAway.

Cancellation policy (Product)

One of the important factors for both hosts and guests to consider is the cancellation policy. Both Airbnb and HomeAway allow the hosts to choose their cancellation policies from a few options they provide. Airbnb has 3 options, while HomeAway has 5 options available for the hosts to choose from.

Table 5 Cancelation policy options provided by Homeaway[14]

Cancellation policy options	Details
No refund	No refund offered for any reason or timeframe
Strict	Bookings cancelled at least 60 days before the start of stay will receive 100% refund
Firm	Bookings cancelled at least 60 days before the start of stay will receive 100% refund. Bookings cancelled at least 30 days before the start of stay will receive a 50% refund
Moderate	Bookings cancelled at least 30 days before the start of stay will receive 100% refund. Bookings cancelled at least 14 days before the start of stay will receive a 50% refund
Relaxed	Bookings cancelled at least 14 days before the start of stay will receive 100% refund. Bookings cancelled at least 7 days before the start of stay will receive a 50% refund

Table 6 Cancellation policy options provided by Airbnb[5]

Cancellation policy options	Details
Flexible	Full refund within a limited period (usually 24 hours before check-in)
Moderate	Full refund within a limited period of time (usually 5 days before check-in)
Strict	Full refund if cancellation is within 48 hours of booking

Just by looking at the number of options, it seems like HomeAway is a more favorable site for the hosts to rent properties. Depending on their financial status and popularity, the hosts get to decide what cancellation policy they want to implement. However, the fact that HomeAway has "No refund" as their option suggests that HomeAway may be friendly for the hosts but not for the guests. Moreover, when comparing both HomeAway and Airbnb, Airbnb has a looser Cancellation policy than HomeAway which may attract more customers. For example, comparing the "Moderate" policy from both HomeAway and Airbnb, HomeAway's "Moderate" policy requires the guest to cancel the booking 30 days before the stay for a 100% refund while Airbnb provides a full refund if cancellation is within 5 days before the stay.

This difference makes Airbnb easier to book a property for the guests then HomeAway because the guests know that they have a greater possibility of getting a refund even if their trip schedule suddenly changed. This also suggests that stricter cancellation policy may tend to make the guests take a longer time to make a decision on reserving a property. Looser cancellation policy may result in a higher risk of getting their rent to be canceled. However, with the guests taking a longer time to make a decision on reserving a

property, the chance of guests deciding not to reserve the place may increase as well. Therefore, Airbnb might be benefiting from the cancellation policies they offer than HomeAway, because by implementing looser cancellation policies, they can attract more customers, followed by increased opportunities to expand the customer base.

Fees (price)

There are two types of fees which are the primary sources of vacation rental businesses' revenue: host fees and service fees. Host fees are fees that are charged from the hosts for each reservation made.[15] Service fees are fees that are charged from the guests for each reservation made.[16]

There is two options host can choose from for HomeAway: Annual subscription and Pay-per-booking.[16] Unlike Airbnb who only offers a pay-per-booking system, HomeAway seems to be more flexible to the different situation hosts are at. However, the host fees are one of the reasons for a number of properties to not increase in HomeAway as it may be discouraging the new hosts to lend properties through HomeAway.

For the host who would want to rent their properties unlimitedly throughout the year for their business, they may choose the annual subscription which is $488 per year.[16] For hosts who only rent their properties out certain times a year, pay-per-booking is recommended. In this case, the host fees will be a minimum of 8% commission on each rental.[16] However, compared to Airbnb, HomeAway's host fees are expensive. Airbnb only uses a pay-per-booking system, but the host fees are generally a 3% commission on each rental.[16] Thus, the hosts are more willing to rent properties through Airbnb than HomeAway in terms of fees because they can keep more profit. In fact, recently, according to the Phocuswright study, the growth rate of Airbnb's property managers listing increased from 7% in 2012 to 47% in 2016.[8]

Service fees are also an important factor to compare as it is one of the factors that decide the price guests need to pay to rent a place. Unlike host fees, Airbnb charges higher service fees from the guests than HomeAway.

At Airbnb, the guests are required to pay the service fee which is around 6% to 12% of the rental cost, depending on the size of the reservation. Data were collected from the top 25 properties that were listed in Kyoto City. Based on the data collected, one-night stay in every property has a service fee of 12.2%.[5] Although 25 properties are a small sample size, with this pattern also seen in the top 25 properties that are listed in the New York City, it can be reasonably argued that many of the properties in Airbnb follow this service fee pattern.

On the other hand, HomeAway typically charges 4% to 9% of the rental cost as a service fee. Similarly, to Airbnb, based on the data collected from the top 25 properties that were listed in the Kyoto city, a one-night stay in every property had the service fee between 5.2% to 10.2%. The average service fee was 8.8%.[5] However, unlike Airbnb, regardless of the length of the stay, the service fee remained almost the same with an average of 8.4%.[5] Again, this pattern was also seen in top 25 properties that was listed in the New York city thus, it can be concluded that HomeAway doesn't have any pattern with its service fees but has an inconsistent service fee of around 8.5% regardless of the length of the stay.

Comparing each service fee, since HomeAway charges a fewer percentage for the service fee, it may attract more customers as it will require fewer payment for them. Partly for this reason, HomeAway had more bookings per year, followed by greater revenue than Airbnb until 2016. Nonetheless, high host fees have contributed to causing a greater gap in the number of properties listed between Airbnb and HomeAway. Thus, Airbnb has now become the market leader creating a greater gap between HomeAway every year.

Promotional strategy (Promotion)

With the vacation rental industry becoming more familiar with the general public, both companies have a great opportunity to expand their customer base. Thus, it is significant for HomeAway to advertise and improve its brand recognition for its future growth. Both HomeAway and Airbnb don't undertake many above-the-line promotions, instead, they both tend to rely on advertisement through social media such as Instagram.[17] However, both Airbnb and HomeAway occasionally carries out a campaign to improve their brand recognition.

One of the HomeAway's recent successful promotion campaigns is called #EiffelTowerAllYours campaign. This was initially planned to celebrate becoming the national sponsor in France at the UEFA EURO 2016.[18] Becoming a national sponsor of France itself gave a huge benefit for HomeAway. In fact, by becoming the national sponsor, HomeAway was able to increase its inventory of rental property listings in France and many of the UEFA fans visiting France stayed in an entire house during the tournament.[19] Moreover, by furthermore conducting the #EiffelTowerAllYours campaign, HomeAway was successfully able to use this opportunity to raise brand recognition.

HomeAway explained that the main aim of the #EiffelTowerAllYours campaign was to "celebrate romantic love, love of family, all kinds of love at the world's most iconic symbol of love".[5] Many people tweeted about this campaign on twitter and HomeAway was able to build a strong word-of-mouth promotion.

This campaign was also reported on many media including the Yahoo! News[20] which also contributed in improving HomeAway's brand recognition. Regarding their successfulness, HomeAway was awarded as one of the winners for the "PRWeek US Awards 2017" under the section for "best global effort".[21] This further improves brand recognition and may give HomeAway a competitive advantage and provide an opportunity for them to increase sales revenue.

However, HomeAway has not been the only one who is successful with the campaigns. Recently, Airbnb has been putting more effort into conducting campaigns. For example, the "We accept" campaign was a campaign about introducing non-discrimination policy on their site since some of their hosts were discriminating against guests based on their race and gender.[8] This has successfully improved the customer experience and reinforced the idea of the global community Airbnb is trying to create. However, one of the advertisement campaigns they carried out caused a conflict between HomeAway, which was called the "Home to You" campaign.[22]

This campaign was about creating a miniature version of the houses that are on their listing from all around the world and hand them on the "Tree of Life" in New Orleans for the migratory birds. This was aimed to convince how "every traveler deserves a home" by using birds and birdhouses a metaphor for its customers and their services.[21] HomeAway sued Airbnb for this marketing campaign as unfair competition. Since 2005, HomeAway has continued to use the logo of the birdhouse, the use of birdhouses for Airbnb's campaign was an action that may confuse the customers. As a result, Airbnb has quietly canceled this campaign. Through this, HomeAway was able to show its users and investors their responsibleness for protecting their trademark which may have contributed to gaining more trust from them.[23]

Since these are just some examples of the campaign each company carries out, it is difficult to determine the level of contribution of these promotional campaigns in creating a gap between HomeAway and Airbnb. However, with Airbnb starting to introduce more below-the-line promotions, this can be a threat to HomeAway. Recently, Airbnb has announced to start a new service called "Loyalty Program" for the frequent hosts. Hosts with ratings of 4.8 or above, at least a 90% response on the ratings, 10 or more stays a year, and zero cancellations are awarded additional payments to the hosts.[24] They furthermore announced that this "Loyalty program" will apply to "loyal customers" as well and they will receive cash back and additional special services.[25] This may be a threat for HomeAway because Airbnb will benefit from this new service since this might motivate both hosts and guests to lend and rent more properties from Airbnb, contributing to their growth.

Conclusion and suggestions

Through this analysis, it is clear that HomeAway has its strengths and weaknesses compared to Airbnb. Their brand recognition has improved by becoming the national sponsor of France and through its successful promotion strategy especially the #EiffelTowerAllYours campaign. However, regarding the rapid growth of Airbnb, in order for Expedia to support HomeAway to increase its market share in the rental vacation industry, there are few things they need to adjust.

First, Expedia must invest more into HomeAway in order to help them diversify the properties they offer like Airbnb, not just the whole property. Although they stated their USP to be lending the whole property, since Airbnb has also started to rent the whole property, it is a threat for the HomeAway to lose its customer base. Also, since HomeAway has a fewer number of property listings, guests are more likely to be in favor of Airbnb because they offer more listings of renting properties to choose from. Furthermore, Airbnb's large number of properties on listings allows people to become more familiar with the brand and strengthen contributes to a further increase in its customers.

As a start for HomeAway to increase its listings can be done by changing their business model and earn profit mainly from the guests, not the hosts. With the current business model, HomeAway relies on profits gained from the hosts then its guests as they charge higher host fee then the service fee. Thus, people who want to rent their properties are more likely to choose to rent at Airbnb which offers cheaper host fees. In order for HomeAway to compete with Airbnb, they should offer cheaper host fees and more expensive guest fees to increase the number of properties. However, with their current listings, since they only rent whole properties, the prices tend to be expensive. Thus, although there is a risk that higher guest fees can limit their customer base, even more, to compete with Airbnb and increase brand recognition, it is important that HomeAway now prioritizes increasing their number of properties on their listings.

Moreover, Expedia needs to plan some strategies to encourage both hosts and guests to use HomeAway. With Airbnb starting a new "loyalty program", Airbnb is likely to start attracting more guests than HomeAway in the near future. Thus, in order to compete with Airbnb, Expedia should consider adjusting the cancellation policies HomeAway offer and make it looser in order to attract more customers. Also, it may also be better for Expedia to consider carrying out more above-the-line promotions for the HomeAway brand. With HomeAway currently having fewer properties on its listings than Airbnb, its brand may be less familiar with the public. Thus, through promoting the brand more, HomeAway can improve its brand recognition to compete with Airbnb and attract more customers.

With Expedia threatened by Airbnb's rapid growth, if Expedia continues to put more effort into their HomeAway brand, they might be able to slow the Airbnb's growth and maintain their position as a market leader in the OTA industry. This may further open up the opportunity for Expedia to keep expanding its business and may also be able to achieve a larger market share overall than Booking Holdings in the future.

References

[1] “Company-Histories.com.” Expedia, Inc. -- Company History, www.company-histories.com/Expedia-Inc-Company-History.html.

[2] “Expedia and Booking Holdings: Revenue 2007-2018 | Statistic.” Statista, www.statista.com/statistics/225455/priceline-total-revenue/.

[3] Schaal, Dennis, et al. “Hotel and Online Travel Agency Direct Booking Winners and Losers in 5 Charts.” Skift, 29 Aug. 2017, skift.com/2017/07/10/hotel-and-online-travel-agency-direct-booking-winners-and-losers-in-5-charts/.
[4] Demmitt, Jacob. “Expedia Completes $3.9 Billion Acquisition of Airbnb Competitor HomeAway.” GeekWire, 16 Dec. 2015, https://www.geekwire.com/2015/expedia-completes-acquisition-of-airbnb-competitor-homeaway/
[5] “Book Your Vacation Rentals: Beach Houses, Cabins, Condos & More.” *HomeAway.com*, www.homeaway.com/.
[6] Bloomberg.com, Bloomberg, www.bloomberg.com/news/articles/2018-08-08/expedia-is-coming-after-airbnb-booking-in-home-rental-market.
[7] “• Airbnb Statistics (2019): Users, Booking, Financials & Fun Facts.” • Airbnb Statistics (2019): Users, Booking, Financials & Fun Facts, 16 Jan. 2019, muchneeded.com/airbnb-statistics/.
[8] VRM Intel Staff, and VRM Intel. “Airbnb vs. HomeAway: Who Is Winning the Race to the Top of the Vacation Rental Industry?” VRM Intel, 27 Dec. 2016, www.vrmintel.com/airbnb-vs-homeaway-winning-race-top-vacation-rental-industry/.
[9] “Vrbo Mission, Vision & Values.” Comparably, HomeAway, www.comparably.com/companies/vrbo/mission.
[10]Sporxx, et al. “Top 100 International Tourist Destination Cities By Country.” Brilliant Maps, 24 Jan. 2015, brilliantmaps.com/top-100-tourist-destinations/.
[11] “MarketMinder: Cutting Edge Vacation Rental Insights.” AirDNA, www.airdna.co/vacation-rental-data.
[12] Blodgett, Sequoia. “Airbnb Aims to Make 'Airbnb for Everyone'.” Black Enterprise, Black Enterprise, 23 Feb. 2018, www.blackenterprise.com/airbnb-aims-to-make-airbnb-for-everyone/.
[13] “2019 Airbnb Statistics - User & Market Growth Data [Updated].” IPropertyManagement.com, ipropertymanagement.com/airbnb-statistics/.
[14] Barker-Pendree, Nick. “HomeAway TV Commercial – It's Your Vacation Why Share It 60s.” YouTube, YouTube, 18 Feb. 2016, www.youtube.com/watch?time_continue=55&v=Y7OeE4-rBkU.
[15] Host an Experience on Airbnb, www.airbnb.com/host/experiences?from_nav=1.
[16] “HomeAway vs Airbnb.” Lodgify, www.lodgify.com/vacation-rental-guide/homeaway/airbnb/.
[17] “7 Magical Branding Lessons from Airbnb.” Laura Busche - Brand Consultant, Author & Speaker, 6 Sept. 2018, laurabusche.com/brand-strategy/airbnb-branding/.
[18] “HomeAway® Makes History by Offering the Eiffel Tower as the Ultimate Holiday 'Rental' : News.” Saatchi & Saatchi, saatchi.com/fr-FR/news/homeaway-makes-history-by-offering-the-eiffel-tower-as-the-ultimate-holiday-rental/.
[19] UEFA.com. “Abritel-HomeAway: EURO's Sixth National Sponsor.” UEFA.com, www.uefa.com/insideuefa/about-uefa/administration/marketing/news/newsid=2322878.html?redirectFromOrg=true
[20] “Here's How You and Five Guests Can Sleep over at the Eiffel Tower This Summer.” Yahoo! News, Yahoo!, 19 May 2016, news.yahoo.com/heres-five-guests-sleep-over-eiffel-tower-summer-161512811.html.
[21] December 08, 2016. “PRWeek US Awards 2017 Shortlist Revealed.” PR Week, www.prweek.com/article/1417913/prweek-us-awards-2017-shortlist-revealed.
[22] “AIR BNB 'Birds'.” Vimeo, 11 May 2019, vimeo.com/125996857.
[23] Santana, Danni, et al. “HomeAway Sues Airbnb Over Birdhouse Trademark Infringement.” Skift, 16 Apr. 2018, skift.com/2013/12/27/homeaway-sues-airbnb-over-birdhouse-trademark-infringement/.
[24] “Airbnb Is Adding a Loyalty Program to Reward You for Booking.” Travel + Leisure, www.travelandleisure.com/travel-news/airbnb-loyalty-program.
[25] Facebook.com/thetripsherpa. “Proposals For An Airbnb Loyalty Program.” Live and Let's Fly, 3 Jan. 2019, liveandletsfly.boardingarea.com/2018/06/24/proposals-for-an-airbnb-loyalty-program/.

Rebranding of General Motors

Kazuma Takahashi
Ritsumeikan Uji High School, Kyoto, Japan
Email: Kazuma.takahashi0103@gmail.com

***Abstract* – General Motors Company, the American multinational automobile company is launching electric bike (e-bike) business from the second quarter of 2019. GM has been unsuccessful in branding as they have sold cars with low MPG in the past and have created negative customer perception towards their product in terms of eco-friendliness. Despite this fact, GM aims to become environmental leader and to stimulate the success of the car industry, GM is entering e-bike business. Therefore, this paper examines the effectiveness of GM's decision to enter e-bike business in terms of rebranding.**

INTRODUCTION

General Motors Company (GM) is an American multinational corporation which designs, manufactures, and distributes vehicle and its components. It is the largest automobile company in U.S., and one of the largest companies in the world in terms of the revenue and its scale of operation. GM have established many different car brands including Chevrolet, Cadillac, and GMC.

Despite its strong revenue over the years, in 2018, there was sliding sales in GM's two largest markets, the U.S. and China with decline in overall unit sale by 1.6%.[1]Moreover, GM Europe was continuously operating at a loss from 1999 to 2016 with loss of over 20 billion dollars and has withdrawn from the European market.[2] GM is preparing for an even greater downturn in the future.

The automobile industry is currently shifting to producing and selling electric vehicles (EV). Electric vehicles are gaining market share with more than 50% of compound annual growth rate over 2012 to 2017.[3] Many automobile companies are attempting to adapt to this changing external environment. GM is following the same path as well, introducing new electric vehicle such as Chevrolet Volt EV. However, EV market share of GM in the U.S. of 2018 is 10% which is far away from the market leader, Tesla which has 53% of EV market share.[4] From this, it is clear how GM is falling behind in the EV market.

One of the factors that is affecting the performance of GM in the EV market is the GM's brand image. In GM's vision statement, it says, "we envision zero emissions and we have committed ourselves to leading the way toward this future" and "we are committed to an all-electric future" which shows how GM is focusing towards establishing environmentally considerate image. However, research from Northwestern University has found out that GM has been making the most vehicles with 15 MPG (miles per gallon) or worst which left GM with limited credibility as an environmental leader.[5] In order to adapt to the new environment and stay competitive, GM needs to change the image that they hold so they need to focus on rebranding as an environmentally friendly company.

The company is expanding into the new market and providing new product which they are using diversification strategy. GM had been investing into electric bicycle (e-bike) business from 2016 because of the rising market share of e-bike. E-bike is a bicycle with integrated electric motor for the ease of cycling. GM has established a new e-bike brand, Ariv. They have designed and manufactured 2 types of e-bike, Merge and Meld, shown in Figure 1 which is an innovative product from its special features.

Figure 1 Ariv Meld and Ariv Merge

These new products are planned to be launched 2nd quarter of 2019 in Europe. The e-bike is said to be a replacement for motorcycle, scooters, and cars and because e-bike does not emit CO2 gas, it is considered as eco-friendly alternative choice. By developing this e-bike business, GM has an opportunity to embrace the eco-friendly image to further improve on their EV market. This means that success in the e-bike business in terms of sales would help GM to establish the eco-friendly and innovative image. Moreover, other automobile companies are expanding into the e-bike business for similar purpose: to embrace the eco-considerate image and to increase their automobile sale. Thus, the question that would be discussed in this paper is "To what extent will GM's decision to enter the e-bike market will be an effective rebranding strategy?". In this paper, the aim is to evaluate the effectiveness of GM's decision in terms of rebranding their company's image. This question will be

[1] ("General Motors")
[2] (*After Losing $20 Billion over 17 Years, General Motors (GM) May Quit Europe by Selling Opel and Vauxhall to Peugeot — Quartz*)
[3] (Ltd)
[4] (*2018 EV and Tesla Market Share Update – S1DD.COM*)
[5] (Ludwiszewski and Haake)

examined through SWOT analysis because it examines the internal and external factors that influence the success of this decision. When changing corporate objectives, this change can be a response to the internal or external factors and GM is shifting its focus and vision on electrification. Therefore, SWOT analysis would be an effective tool to analyze GM's changing objectives and decision. Then the example of the success of Apple Inc's rebranding would be examined to connect to GM's case. This would help evaluate GM's rebranding strategy by comparing with a company that have succeeded in rebranding.

ANALYSIS AND DISCUSSION

Strength

GM is a 4th largest revenue generating automobile company in the world with 147 billion dollars in 2018, they can benefit from the economies of scale.[6] They can make their cost of production lower because GM supplies their materials from the bulk to get advantages. From the large finance that GM has, they can benefit from the purchasing economies and manufacture their e-bike with low cost.

Another strength is that GM is using above the line promotion which can attract wide range of customers. GM has an enough financial power to use the expensive above the line promotions such as commercial videos which is shown in figure 2.[7] Also, they use these promotions in online article and YouTube which they have diverse promotional mix. This is a strength for GM because there would be a higher chance that the new product would be

Figure 2 Commercial of Ariv

recognized by more people. This strength can foster their brand image as environmentally friendly company.

Another strength is that they have experience and human resource for producing EV motors. This is because GM is currently focusing on electrification in their automobile company, producing electric cars which involves in creating EV motors. GM has been investing in the electric motor from 2011 having a vision to establish GM as a leader in automotive electric motors.[8] From 2013, GM has started producing electric motor needed for manufacturing electric vehicles by constructing the electric motor plant. GM was the first company to make electric motor from all the major US automobile company. This indicates the strong experience and knowledge for the electric motor which is also required to manufacture e-bikes. From this strength, GM was able to install a lithium ion battery which has largest capacity while having lightest weight out of all the other battery types. This would be a strength for GM to more effectively manufacture and produce the electric bike which would further embrace the image as environmentally friendly company.

Lastly, the innovative feature of the e-bike that they are producing is another strength. GM is currently planning to offer has a special feature which is that the e-bike can be folded compactly.[9] This feature is not implemented in other e-bike company that is a market leader in Europe where most of them focus on sports type e-bike. This means that by creating innovative product, they were able to differentiate with the other company that are functioning in the European market. Moreover, there will be less competition in this segment which would be a GM's strength to sell folded e-bike as main product in GM's e-bike business.

Also, as an innovative strategy, GM's e-bike can be connected with the smartphone app of GM which gives various information to the riders such as the distance, speed, and remaining battery.[10] This special feature can be a strength for GM as it can be a USP of the company for differentiating between other company's product. Integration of other technology would lead to further promoting the company and emphasizing on the fact that GM is an environmental caring company.

 [11]

Figure 3 Integration with App

Weakness

One of the weakness would be the delay of GM entering e-bike business as, there are already many companies in this market. A lot of the companies such as Yamaha and SRAM and Shimano have already functioning in the market way before GM entered the market. For example, Yamaha has launched its first e-bike in April of 2018.[12] Since innovation is greatly valued in business because business need to be up to date, the delay for entering the market can be a weakness. Moreover, GM has no experience in the bicycle or e-bike company because they were consistently involving in the automobile market. Therefore, GM has limited brand recognition as company selling e-bike, so they need to effectively promote their new business and product in order to succeed in e-bike business

[6] ("Focus2move| World Car Group Ranking - The Top 25 in the 2019")
[7] (*ARĪV Merge E-Bike 2019*)
[8](*GM to Build Its Own Electric Motors | WIRED*)
[9] (*ARĪV Electric Bike | ARĪV Netherlands*)
[10](*ARĪV Merge E-Bike 2019*)
[11] (*ARĪV Electric Bike | ARĪV Netherlands*)
[12] (Siler)

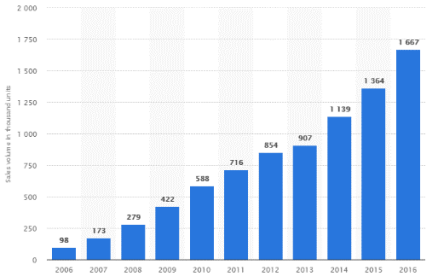
and embrace the image as an environmentally friendly company.

Another weakness is the fact that GM's automobile business has experienced withdraw from European market as they have consistently operated in negative profit.[13] This is a weakness as the poor profitability and sales of GM product in the past would give a negative image towards GM in this region. Customer loyalty and trust is an essential part in building customer base so this would be a weakness for GM.

Also, GM had a massive scandal in 2014 where they needed to recall 2.6 vehicles with ignition switch defect.[14] The troubles in the vehicles were from the mistake in the manufacturing process. GM needed to pay a fine of 900 million dollars and needed to comply with the issue for 3 years.[15] This issue had an impact globally where the defected vehicles were distributed worldwide including European market. This long-term scandal gave negative brand image towards the company which lost many trusts and loyalty from the customers and stakeholders which would be a weakness in promoting and selling the new product.

Opportunity

As an opportunity, GM's Ariv E-bike is operating in Europe which is the area where there is highest demand for e-bike in the world. E-bike business in Europe has shown more than 15% growth in compound annual growth in 2017.



16

Figure 4 Sale Volume of E-bike (in thousands) in Europe

As figure 9 shows, there are steady growth in the sales of e-bike in European market. This indicates the opportunity for GM for them to use the external environment as their opportunity to raise their sales of e-bike. GM's decision to operate the new e-bike business in Europe may be an effective decision as it is the region that has the highest demand and is the highest revenue generating region.

Specifically, GM is operating in 3 countries, which are Belgium, Netherlands, and Germany. In Germany, there are annual record growth in the sale of e-bike. From 2016 to 2017, there was 19% increase in e-bike sales[17] and from 2017 to 2018, there was 39% increase in sales with one million units sold.[18] For Netherland, the unit of e-bike sold is greater than unit of regular bicycle sold indicating high demand for e-bike.[19] Also, the e-bike market in Netherland is growing in a steady pace with 8.7% increase in sale volume in 2018.[20] In Belgium, the sales of e-bike is also growing continuously with the support of the tax incentive of the government.[21] Out of all the sales of bicycle in Belgium, e-bike accounts for 45% which is the highest percentage indicating a strong opportunity.("E-Bikes Take Lead in Belgian Market")[22] Overall, the countries and region that GM is operating has a growing e-bike market, which is a large opportunity for GM to raise the sale of e-bike in the market and contribute to the rebranding.

Another opportunity is the worldwide movement towards climate change and society's consciousness towards the issue. From the political factor in external factors, there are possibility of government grants for GM since they are producing e-bikes which is considered as environmentally friendly choice. Not only by the aim to reduce CO2 emission, the government wants to emphasize on the health as an alternative for cars. Like this, e-bike has many characteristics which can be supported by the government to give a further advancement in this business.

Lastly, the cost for producing lithium ion battery is decreasing which is used by GM.[23] According to Boomerang NEF, the average annual decline in the cost of lithium ion battery from 2010 to 2018 was 20.5%.[24] It is forecasted to decrease 53% from 2018 to 2024 which shows the opportunity for GM to reduce the cost to produce the product in the future. The strength of the battery types of GM's e-bike can be used to maximize the opportunity of the declining cost for lithium ion battery. From this, GM has a greater chance to succeed in the market and thus has an opportunity to distribute more e-bike in Europe to embrace the environmentally friendly image.

Threats

Firstly, the diversification strategy can be a threat for GM as they have a high competition between other companies which already has their customer base. GM motor is using diversification strategy as they are planning to make new product to the new customers as they have been previously focused on only making a vehicle or vehicle

[13] (*GM to Be First U.S. Automaker to Make Electric Motors*)
[14] (Press)
[15] (Press)
[16] ("Electric Bicycle Sales EU 2006-2016 | Statistic")
[17] ("German E-Bike Market Sees Record Growth in 2017")
[18] ("One Million E-Bikes Sold in Germany in 2018; Up 36 Percent!")
[19] ("E-Bike Puts Dutch Market Back on Growth Track")
[20] ("E-Bike Puts Dutch Market Back on Growth Track")
[21] ("E-Bikes Take Lead in Belgian Market")
[22] ("E-Bikes Take Lead in Belgian Market")
[23] (*GM Releases ARĪV Electric Bicycle with Two Models Availble to Pre-Order*)
[24] ("A Behind the Scenes Take on Lithium-Ion Battery Prices")

parts. Out of four strategies presented in the Ansoff Matrix, diversification has the highest level of risk as GM needs to expand to market with less experience. In addition, they are entering a new market which they would have to deal with new customers. On the other hand, the typical e-bike company such as Bosch, Yamaha and SRAM have been involved in the bicycle industry previously before they have expanded their business to e-bike.[25] They have been manufacturing and selling standard bicycle and they have developed their bike to use electric power, so this is product development strategy. These companies have already established their brand image as a bicycle company which would increase their reliability to their product and would generate customer loyalty. It would be a threat for GM that they need to compete with the company that already have established customers for the bicycle market. The weakness that GM is a company entered e-bike market more late than other company has affected the threats that GM might have disadvantages over other competitors.

Also, because GM has chosen to implement e-city bikes, it would make the competition much intense. 99% share of the e-bike sold in Germany is 250Watt, 25km/h models and out of that, the highest share is 38.5% which is the sale of e-city bikes.[26] This means that there are high level of competition in this specific e-bike market so it would be a threat for GM. However, it can be an opportunity as well since it indicates the high demand from the European customers and the possibility of succeeding in sales of the product.

Lastly, there is a possibility that the national regulations might be posed to limit the use of e-bike in Europe. In 2016, it was reported that China has banned the use of electric two-wheelers including e-bike in several major city in China such as Shenzhen, Beijing and Guangzhou.[27] The reason for the regulations was that there were extremely large number of people who use e-bike in China which were usually violating the traffic law. Study by Christopher Cherry, a professor and researcher of transportation in China showed that e-bike users were driving in the wrong direction on 44% and violation stop signs and traffic signals for 56%. In fact, the e-bike fatalities in China are high which is increasing in an exponential rate.[28] The Chinese government couldn't have enough traffic control which they needed to ban the use of e-bike. This is a serious threat for GM because there can be a regulation posed if Europe has a same situation. The opportunity of the growing e-bike user in Europe can be a threat by a possibility of overflowing e-bike users in Europe which might cause regulations for the use of e-bike in this region. Europe is the only market segment that GM is considering in the current point and if regulations are posed, it would limit the sales of e-bike in Europe. This would further affect the brand image that GM is attempting to embrace because of the lack of sales which is ineffective for recognition.

Rebranding

Rebranding is a market strategy which is intended to develop new brand identity to better attract customers. Branding is an important factor to establish corporate identity: where business want to place themselves in the market. Business often rebrands for their own purpose. According to the study of 165 cases of rebranding by University College Dublin, there are different purpose when business rebrands, and there are several ways they can do so.[29] There are three types of rebranding which are corporate rebranding, business unit rebranding, and product rebranding.[30] GM is aiming for corporate rebranding as they are focusing on the rebranding of the company overall.

Also, there are four driving forces of rebranding which are change in corporate structure, change in corporate strategy, change in competitive position, and change in external environment.[31] One of the main reasons why business use rebranding strategy is to alter customer's negative perception of the company.[32] In the case of GM, the customer perception is shifting to more compact cars with less pollution. Also, the EV markets are extremely becoming popular. From these external changes, GM needed to embrace eco-friendly image. To grow the sales of GM's EV which is GM's main aim, GM has diversified into e-bike business.

As an example of successful rebranding, Apple Inc. can be considered. Apple was originally only producing computers. Since the sales of their computer was decreasing, Apple has diversified into smartphone business to increase their personal computer sales.[33] By using sub-products, they aimed to embrace innovative brand image to increase the sales of their main product. When Apple entered the smartphone market in 2007, other companies were already in the market.[34] However, Apple's smartphone had an innovative feature such as the compact body, no keypads, and long battery life[35] and became a unique selling proposition (USP) which differentiated from the smartphone produced by other companies. Apple succeeded in rebranding them as innovative company by using sub products that had a rising market growth. We can see how innovation is a key to the rebranding. This is similar to GM's situation of how GM is trying to increase their car sales by investing into the growing market of e-bike to rebrand their image. GM needs to rebrand themselves as sustainable: ecofriendly, and a company that has high

[25] ("German E-Bike Market Sees Record Growth in 2017")
[26] ("German E-Bike Market Sees Record Growth in 2017")
[27] (Zuev et al.)
[28] (Shepard)
[29] (Laurent Muzellec)
[30] (Laurent Muzellec)
[31] (Laurent Muzellec)
[32] (Laurent Muzellec)
[33] ("How Apple Changed the World - 4 Core Marketing Strategies of the Tech Icon")
[34] (*The History and Evolution of the Smartphone*)
[35] ("How Apple Changed the World - 4 Core Marketing Strategies of the Tech Icon")

technology to produce good quality EVs. As a method, by producing innovative e-bike, they can differentiate from other e-bike companies for rebranding of GM to eventually increase the sale of their automobiles and EVs. As stated in the strength in the SWOT above, GM's e-bike has innovative features such as integration with technologies. This indicates how GM's innovation in e-bike is vital in the process of rebranding, thus GM's decision to expand into the e-bike business is effective for their rebranding in this sense.

CONCLUSION

From the discussion above, entering the e-bike business for GM to effective brand image to a great extent. It was clear from the analysis that GM has a financial strength to obtain materials, and manufacture, and promote their e-bike effectively. GM was able to use diverse promotional strategy and take advantages of the economies of scale. Additionally, GM has strong human resource with knowledge and experience of manufacturing and producing electric motor which is a key component of producing e-bike. From their strong human resource and strong developing team, GM was able to produce innovate features such as bicycle frame that can be folded and integration system with smartphone apps that are installed in GM's e-bike. Furthermore, GM has a large opportunity to as they have large chance of success by the growing market of e-bike in Europe, which is increasing in an exponential way. Also, it was identified that there is a global trend of governments attempting to cope with climate change issues, which can be a great opportunity for GM to receive governmental grants. This opportunity from political factors was derived from the GM's strategy which is to focus on electrification instead of using energy source that emits CO2 gas. The cost of the lithium battery which is used in GM's e-bike is decreasing over time which shows the chance of GM becoming more profitable in the future. These chance of success in e-bike business would connect to foster the environmentally friendly image that GM is attempting to establish. However, GM's entrance in e-bike market had delayed compared to other major e-bike companies operating in Europe. Also, GM had a scandal and withdraw from European market which created a negative image to the company as a whole. These weaknesses create threats such as intense competition in the market and customer not having a good image towards the new business because of the negative image. We have also found out that there might be a national regulation made like what happed to China. However, competition is not negligible for all businesses which can cause stimulation for GM and motivations for the GM's employees. Moreover, despite the possibility regulations in countries in Europe, GM has countless strength and opportunities which outlaws these negative factors. GM needs to keep in mind of all of the disadvantages and threat but by overall, GM has a great potential to succeed in their e-bike business which would aid them to establish environmentally friendly image to even give a further success in their automobile business as well. Considering the case of Apple, it was further identified how GM's innovation in e-bike is effective in the process of rebranding. Therefore, it is reasonable to say that they should proceed in manufacturing and promoting their new e-bike business.

REFERENCES

2018 EV and Tesla Market Share Update – S1DD.COM. https://s1dd.com/2019/01/2018-ev-and-tesla-market-share-update/. Accessed 26 June 2019.

"A Behind the Scenes Take on Lithium-Ion Battery Prices." *BloombergNEF*, 5 Mar. 2019, https://about.bnef.com/blog/behind-scenes-take-lithium-ion-battery-prices/.

After Losing $20 Billion over 17 Years, General Motors (GM) May Quit Europe by Selling Opel and Vauxhall to Peugeot — Quartz. https://qz.com/914573/after-losing-20-billion-over-17-years-general-motors-gm-may-quit-europe-by-selling-opel-and-vauxhall-to-peugeot/. Accessed 26 June 2019.

ARĪV Electric Bike | ARĪV Netherlands. https://www.arivmobility.com/nl/en/. Accessed 19 June 2019.

ARĪV Merge E-Bike 2019. 2019. *YouTube*, https://www.youtube.com/watch?v=WCD9Q_Y4WIA.

"E-Bike Puts Dutch Market Back on Growth Track." *Bike Europe*, https://www.bike-eu.com/sales-trends/nieuws/2018/03/e-bike-puts-dutch-market-back-on-growth-track-10133083. Accessed 19 June 2019.

"E-Bikes Take Lead in Belgian Market." *Bike Europe*, https://www.bike-eu.com/sales-trends/nieuws/2018/01/e-bikes-take-lead-in-belgian-market-10132617. Accessed 19 June 2019.

"Electric Bicycle Sales EU 2006-2016 | Statistic." *Statista*, https://www.statista.com/statistics/397765/electric-bicycle-sales-in-the-european-union-eu/. Accessed 19 June 2019.

"General Motors." *Fortune*, http://fortune.com/fortune500/2019/general-motors/. Accessed 25 June 2019.

"German E-Bike Market Sees Record Growth in 2017." *Bike Europe*, https://www.bike-eu.com/home/nieuws/2018/08/german-e-bike-market-sees-record-growth-in-2017-2-10134349. Accessed 19 June 2019.

GM Releases ARĪV Electric Bicycle with Two Models Availble to Pre-Order. https://electrek.co/2019/02/14/gm-ariv-electric-bicycle/. Accessed 25 June 2019.

Rebranding of General Motors

GM to Be First U.S. Automaker to Make Electric Motors. /Read/22433/GM-manufacture-electric-motors. Accessed 19 June 2019.

GM to Build Its Own Electric Motors | WIRED. https://www.wired.com/2010/01/gm-build-electric-motors/. Accessed 19 June 2019.

"How Apple Changed the World - 4 Core Marketing Strategies of the Tech Icon." *Neil Patel*, 12 Oct. 2011, https://neilpatel.com/blog/how-apple-changed-the-world/.

Laurent Muzellec. *Corporate Rebranding: Impact on Brand Equity*. https://www.slideshare.net/LaurentMuzellec/corporate-rebranding-transfr.

Ltd, Motionlab Marketing. "Chevrolet's Withdraw from Europe: Its Impact on Opel/Vauxhall." *JATO*, 9 May 2016, https://www.jato.com/chevrolets-withdraw-europe-impact-opelvauxhall/.

Ludwiszewski, Raymond B., and Charles H. Haake. "Cars, Carbon, and Climate Change." *Northwestern University Law Review*, vol. 102, 2008, p. 665.

"One Million E-Bikes Sold in Germany in 2018; Up 36 Percent!" *Bike Europe*, https://www.bike-eu.com/sales-trends/nieuws/2019/03/one-million-e-bikes-sold-in-germany-in-2018-up-36-percent-10135573. Accessed 19 June 2019.

Press, Associated. "Key Events in General Motors' Ignition Switch Recall." *Inc.Com*, 17 Sept. 2015, https://www.inc.com/associated-press/key-events-in-general-motors-ignition-switch-recall.html.

Shepard, Wade. "Why Chinese Cities Are Banning The Biggest Adoption Of Green Transportation In History." *Forbes*, https://www.forbes.com/sites/wadeshepard/2016/05/18/as-china-chokes-on-smog-the-biggest-adoption-of-green-transportation-in-history-is-being-banned/. Accessed 25 June 2019.

Siler, Ty Brookhart and Wes. "This Is What Happens When a Motorcycle Company Makes E-Bikes." *Outside Online*, 3 Apr. 2018, https://www.outsideonline.com/2293446/what-if-motorcycle-company-made-electric-bicycles.

The History and Evolution of the Smartphone: 1992-2018. https://www.textrequest.com/blog/history-evolution-smartphone/. Accessed 30 Aug. 2019.

Zuev, Dennis, et al. "Where Is the Politics? E-Bike Mobility in Urban China and Civilizational Government11We Are Grateful to the UK's Economic and Social Research Council (ESRC) for Funding This Project (ES/K006002/1), 2013-17." *Environmental Innovation and Societal Transitions*, vol. 30, Mar. 2019, pp. 19–32. *ScienceDirect*, doi:10.1016/j.eist.2018.07.002.

Understanding The Development of Mathematics From The History of The Basel Problem

Jaewoo Lee
Ritsumeikan Uji Junior and Senior High School, Kyoto, Japan
Email: leejaewoo1230@gmail.com

Abstract – **In the field of mathematics, development is still ongoing nowadays. People are constantly finding new theories or formulas to develop the knowledge of mathematics. Currently, there are numerous amounts of mathematical questions that are not yet solved. The Basel Problem used to be one of such questions that many prestigious mathematicians had struggled. Leonhard Euler is the first mathematicians who managed to solve the problem. In the method he used, he mainly used the concept of integrals and trigonometry. In the process of historical development inside mathematics, previous knowledge is always used to explore new knowledge. In fact, Euler only used past knowledge to solve the Basel problem. By analyzing how Euler approached the Basel Problem, I will find out how reliant mathematics is on already discovered concepts. Nowadays, the Basel Problem is applied in other topics. One of the millennium prize problems, the Riemann Hypothesis, is not yet solved, but the solution of the Basel problem can be used as a hint. This represents the mechanism of history in mathematics, that past knowledge is being applying to new knowledge to explore the unsolved pathway.**

Key Words – Basel Problem, Integration, Leonhard Euler, Math, Riemann Hypothesis,

INTRODUCTION OF BASEL PROBLEM

Basel problem, originally proposed by Pietro Mengoli in 1644, asks for the summation of the reciprocals of the squares of the natural numbers, expressed as following [1]:

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

Basel problem can be solved by various methods including, but not limited to, calculus, probability, and complex analysis. In this paper, I will be demonstrating 2 different solutions utilizing Maclaurin expansion and calculus.

These various methods were found during the 18~19th century. In addition to that, the history of this problem critically discussed the significance of different area of mathematics collaborating to other areas of study [1]. For example, there is the unsolved problem called Riemann hypothesis which, in a simple explanation, discuss about the distribution of prime number [2]. In that problem, function called Riemann Zeta function is significantly involved and Basel problem is used in the part of that function. This essay will discuss about the history of Basel problem by focusing at what happened until Euler reached the solution of it.

HISTORY OF BASEL PROBLEM

Basel problem was initially raised by Italian mathematician Pietro Mengoli in 1644 [1]. Although his work is known to few people, he is the first one who stated the problem [3]. The difficulty of Basel problem was harsh in a large extent, even the proponent couldn't manage to solve it [3]. That can be known by the fact that Gottfried Wilhelm Leibniz, known to be the inventor of calculus and master of many infinite series, had challenged and failed to prove [3].

Basel problem gained its fame when Jakob Bernoulli, tutor of Paul Euler -the father and tutor of Leonhard Euler- wrote about his attempt in 1689 [1] [3].

By looking at the Jakob Bernoulli's achievements, it is obvious that he loved the infinite series and he enjoyed solving using p-series [3] [4]. From there he earned the equation below [3]:

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

It is noticeable that if p=2 it will become Basel Problem. Despite he found an equation, he thought he cannot solve it by himself [3]. Instead, he decided to find the estimated value of it [3].

Jakob Bernoulli first started his investigation by exploiting the inequality [3]:

$$2n^2 \geq n(n+1)$$

From the equation he got above, he proved that Basel Problem solution will be smaller than 2 [3].

$$\frac{1}{2n^2} \leq \frac{1}{n(n+1)}$$
$$\frac{1}{n^2} \leq \frac{2}{n(n+1)}$$
$$\therefore \sum_{n=1}^{\infty} \frac{1}{n^2} \leq \sum_{n=1}^{\infty} \frac{2}{n(n+1)}$$

Jakob Bernoulli found what is $\sum_{n=1}^{\infty} \frac{2}{n(n+1)}$ equals to [3].

$$\sum_{n=1}^{\infty} \frac{2}{n(n+1)} = \sum_{n=1}^{\infty} 2\frac{1}{n(n+1)} = \lim_{N\to\infty} 2\left(1 - \frac{1}{N+1}\right)$$

Since N will be approach to infinite, $\frac{1}{N+1}$ will be approach to very small value which I can assume as 0.

$$\sum_{n=1}^{\infty} \frac{2}{n(n+1)} = 2$$
$$\therefore \sum_{n=1}^{\infty} \frac{1}{n^2} \leq 2$$

This equation indicates the fact that solution of the Basel Problem will be smaller than 2.

Even though Jakob Bernoulli's convergence was impressive, this didn't bring him to find the exact value for Basel Problem. As he was publishing his proof of convergence, Jakob Bernoulli requested the mathematicians for help [5]. This Jacob Bernoulli's work showed the example of significance of steps in the development of mathematics. Even though he couldn't reach the solution for Basel problem, assuming based on mathematical proof is significant steps to reach the solution. Because of these piles of invention, Euler was able to face Basel problem and confirm the solution easier, since solution of Basel problem is already known that it will be less than 2.

EULER'S APPROACH TO THE BASEL PROBLEM

Euler first managed to solve the problem in 1729 and 1730 [1]. In this proof, he mainly used Taylor expansion of sine. Despite this proof reached to correct solution, other mathematicians didn't agree with his methodology. Euler, then, published another proof in 1741 which made everyone agree [1]. Interestingly, this proof method only uses fundamental level of calculus, Taylor series, and integrations. It is longer than first proof but easier.

Euler's first step was to use the equation $\frac{1}{2}(sin^{-1}x)^2$ because he already predicted that this equation will lead him to find the Basel Problem.

$$\frac{1}{2}(sin^{-1}x)^2 = \int_0^x \frac{sin^{-1}t}{\sqrt{1-t^2}}dt$$

Euler first differentiated $\frac{1}{2}(sin^{-1}x)^2$ and integrated that equation.

For next equation, Euler focused on the $sin^{-1}x$.

$$g(x) = sin^{-1}(x)$$
$$g'(x) = \frac{1}{\sqrt{1-x^2}} = (1-t^2)^{-\frac{1}{2}}$$
$$\therefore sin^{-1}(x) = \int_0^x (1-t^2)^{-\frac{1}{2}}dt$$

Next, Euler focused on the $\int_0^x (1-t^2)^{-\frac{1}{2}}dt$. By using the Maclaurin expansion, Euler expanded the equation.

$$\int_0^x (1-t^2)^{-\frac{1}{2}}dt = x + \frac{1}{2\times 3}x^3 + \frac{1\times 3}{2^2\times 2!\times 5}x^5 + \frac{1\times 3\times 5}{2^3\times 3!\times 7}x^7 + \cdots$$

After the expansion, Euler started to use previous equations to finish his proof. Starting from the equation he found first:

$$\frac{1}{2}(sin^{-1}x)^2 = \int_0^x \frac{sin^{-1}t}{\sqrt{1-t^2}}dt$$

Let $x = 1$

$$\frac{\pi^2}{8} = \int_0^x \frac{sin^{-1}t}{\sqrt{1-t^2}}dt$$

Since $sin^{-1}(x)$ was equalized to $\int_0^x (1-t^2)^{-\frac{1}{2}}dt$ and expansion was already shown, expand the $sin^{-1}t$ in equation above.

$$\int_0^1 \frac{sin^{-1}t}{\sqrt{1-t^2}}dt = \int_0^1 \frac{t + \frac{1}{2\times 3}t^3 + \frac{1\times 3}{2^2\times 2!\times 5}t^5 + \frac{1\times 3\times 5}{2^3\times 3!\times 7}t^7\cdots}{\sqrt{1-t^2}}dt$$
$$\int_0^1 \frac{sin^{-1}t}{\sqrt{1-t^2}}dt = \int_0^1 \frac{t}{\sqrt{1-t^2}}dt + \frac{1}{2\times 3}\int_0^1 \frac{t^3}{\sqrt{1-t^2}}dt + \frac{1\times 3}{2\times 4\times 5}\int_0^1 \frac{t^5}{\sqrt{1-t^2}}dt + \cdots$$

Since $\int_0^1 \frac{sin^{-1}t}{\sqrt{1-t^2}}dt = \frac{\pi^2}{8}$

$$\frac{\pi^2}{8} = \int_0^1 \frac{t}{\sqrt{1-t^2}}dt + \frac{1}{2\times 3}\int_0^1 \frac{t^3}{\sqrt{1-t^2}}dt + \frac{1\times 3}{2\times 4\times 5}\int_0^1 \frac{t^5}{\sqrt{1-t^2}}dt + \cdots$$
$$\frac{\pi^2}{8} = 1 + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \cdots$$
$$\frac{\pi^2}{8} = \frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots$$

Aside from this equation, Euler built another equation, which he expanded the first few terms of Basel Problem.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots\right) + \left(\frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \frac{1}{8^2} + \cdots\right)$$

Euler divided the terms with odd number and even numbers.

Since $\frac{\pi^2}{8} = \frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots$,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{8} + \left(\frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \frac{1}{8^2} + \cdots\right)$$

For the term $\left(\frac{1}{2^2}+\frac{1}{4^2}+\frac{1}{6^2}+\frac{1}{8^2}+\cdots\right)$, Euler changed its form.

$$\left(\frac{1}{2^2}+\frac{1}{4^2}+\frac{1}{6^2}+\frac{1}{8^2}+\cdots\right)=\left(\frac{1}{4}+\frac{1}{16}+\frac{1}{36}+\frac{1}{64}+\cdots\right)$$
$$=\frac{1}{4}\left(1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\right)$$
$$=\frac{1}{4}\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots\right)$$

Notice that at this point, another "Basel Problem" is appearing in the parenthesis.

$$\left(\frac{1}{2^2}+\frac{1}{4^2}+\frac{1}{6^2}+\frac{1}{8^2}+\cdots\right)=\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{n^2}$$

It was already shown that:

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{8}+\left(\frac{1}{2^2}+\frac{1}{4^2}+\frac{1}{6^2}+\frac{1}{8^2}+\cdots\right)$$

Notice that I can substitute previous value:

$$\therefore\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{8}+\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{n^2}$$
$$\sum_{n=1}^{\infty}\frac{1}{n^2}-\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{8}$$
$$\frac{3}{4}\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{8}$$
$$\therefore\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$$

As written in the previous section, Jakob Bernoulli estimated that answer of the Basel Problem will smaller than two. In fact, assuming $\pi = 3.14$ it is possible to examine how much Jakob Bernoulli was accurate with his proof.

$3.14^2 = 9.8596$

$9.8596 \div 6 = 1.64327$

Since $1.64327 \leq 2$, Jakob Bernoulli was reasonably correct with his proof.

RELATIONSHIPS WITH NOWADAYS

After the problems were solved in various ways, mathematicians started to implement the Basel Problem in various topics in mathematics. This is similar to when other areas of mathematics were used to solve Basel Problem. Basel Problem has become the topic which assist mathematicians to solve other problems or explore to new areas of study. This flow displays the process of how the history of mathematics is develoing. in history of mathematics, which shows how the problem is solved and become one of the independent knowledges also collaborating with other sets of knowledge to assist the develop of history of mathematics. In this section, one of the examples of how Basel Problem is used after it was solved will be presented.

Riemann Zeta function is function that is still researched to develop the number theory for solving the prime numbers [6]. In 1737, Euler found the Zeta function, but didn't develop as much as Bernhard Riemann did [6]. Even though it has been long years since it found, it is still researched nowadays. The function is stated as:

$$\zeta(x)=\sum_{n=1}^{\infty}\frac{1}{n^x}$$

Notice that when $x = 2$, it will become the Basel Problem. Solving Basel Problem allowed mathematicians to earn one of the values in Riemann Zeta function. This fact highlights, how the old invention is contributing to the topic established later on and encouraging to the development of mathematics.

Similar to Basel Problem's becoming a part of Riemann Zeta Function, Riemann Zeta Function is the part of Riemann Hypothesis. Riemann Hypothesis is one of the famous unsolved mathematics problems [2]. Riemann Hypothesis is the hypothesis, proposed via paper in "*Monatsberichte der Berliner Akademie*" by Bernhard Riemann in 1859 [2]. In short, by solving this problem, people will be able to find the distribution of prime number. Prime number is used in various cryptograph methods nowadays technology; for example, it has become an important factor to protect personal information. Riemann Hypothesis is involving in a great extent in a current technology and Basel Problem is being the part of key to solve that hypothesis.

Basel Problem was the problem solved a long ago. However, it has been a part of mathematical knowledge and contributing to solve the current hypothesis. By looking at the application of the Basel Problem, it was highlighted that solving Basel Problem was significant in the history of the development of the number theory.

CONCLUSION

In conclusion, by researching the history of the Basel Problem, I was able to observe the process of development in mathematics. The Basel Problem took decades to find a solution even though many mathematicians were involved. Because of its difficulty, after it got solved, it has become an important factor to solve another hypothesis. The Basel Problem's history represents the development in history of mathematics.

When Basel Problem was first stated in 1644, many of the mathematicians tried to reach the solution. Although they were unsuccessful with solving the Basel Problem, they slowly invented steps to reach the solution. Euler, who first solved the Basel Problem, could not have solved or managed all of the question by himself. In fact, he used other mathematicians' clues or theories to reach the solution. That being said, Euler was the one who calculated everything, but other mathematicians' presences were necessary.

The story of solving the Basel problem shows that the factors required to develop the history of mathematics is not only knowledge. It demands other mathematicians' opinions, hypothesis, and theories. Solutions to mathematical problems become part of the knowledge and applied to other parts of mathematics only if the problem is proved and approved by mathematics community. For example, in the process of

solving the Basel Problem, Euler used the Maclaurin expansion. The Maclaurin expansion was not a part of knowledge in mathematics from the beginning. It was found by Colin Maclaurin and slowly became a part of topics within mathematics [8]. This can be applied to the Basel Problem as well. The Basel Problem got solved in many ways, but nowadays, it has become one independent knowledge, providing assistance in the early stage of solving the unsolved number theory hypothesis –Riemann Hypothesis. The long process of continuous discovery and application of ideas has constructed the history of mathematics that we know today and will continue to do so even in the future.

ACKNOWLEDGMENTS

The author wishes to thank Mr. Elliot Lui for his help and guidance in this paper.

REFERENCES

[1] Sandifer, E, C, " Euler's solution of the Basel problem—the longer story," *Journal,* pp. 1-6

[2] Hosch, W, L, "Riemann Hypothesis," *Online* Available at: https://www.britannica.com/science/Riemann-hypothesis [Accessed 6 December 2018]

[3] Dunham, W, "Euler: The Master of Us All," *Book*, Volume 22.

[4] Weisstein, E, W, "p-Series," *Online* Available at: http://mathworld.wolfram.com/p-Series.html

[5] Jakob, B, "Tractatus de seriebus infinitis (Treatise on infinite series)," *Journal*.

[6] Britannica, The Editors of Encyclopaedia, "Riemann Zeta Function," *Online* Available at: www.britannica.com/science/Riemann-zeta-function

[7] Britannica, The Editors of Encyclopaedia, "Taylor Series," *Online* Available at: www.britannica.com/science/Taylor-series

Deriving a Function Which Expresses the Velocity of an Object Rolling Down a Complex Surface

Kaito Nakamichi
Ritsumeikan Uji High School, Kyoto, Japan
Email: nakamichi.kaito29@gmail.com

***Abstract* – This paper presents the possibility of a mathematical formula that expresses the velocity of an object rolling along a complex surface. This function was derived by considering the forces applied to an object. In addition, the change in angle between the slope at a given point on the complex surface and the line parallel to the ground in a 2-dimensional direction will be considered. This mathematical expression may allow increasingly accurate calculations of the theoretical velocity of an object at a given point on a complex surface. Furthermore, it allows a different perspective on how the velocity of an object in this situation can be calculated. However, the expression is not completely accurate and can only be applied under certain assumptions and conditions.**

Key Words - Mathematical derivation, Physics, Mathematics, Velocity of an Object, Formula derivation

INTRODUCTION

Bikers face many challenges when trying to gauge their speed on steep terrains. This can be extremely dangerous since it makes the bicycle harder to manoeuvre. It was hypothesized that understanding the relationship between the speed of the bike and the steepness of the slope could contribute to giving cyclists a better idea of the what speed would they will likely to travel in a specific terrain. None of the expressions or formulas accessible was able to take into account the friction caused by the contact of the rolling object and the surface it is rolling down. Therefore, a method to derive the function which expresses velocity of an object at a given point rolling down a complex surface with minimal information will be explored in this paper.

An attempt was made to derive this equation using directional derivatives and partial differentiation. Finding the gradient of the slope and then deriving an equation for the change in the angle of that slope in order to derive an acceleration function. However, in the process of writing a paper on this methodology, it was recognized that directional derivatives were not able to take the smallest detail of the change in gradient of the function in a given direction into consideration. Instead, directional derivatives only gave an overall or the average gradient of the surface in the direction of a certain vector.

Another method was to model the change in height of the surface to the x, y plane as a 2-dimensional function. This method allows the small details in the change in height of the function. Thus this paper will first examine 3-dimensional surfaces and their connections to 2-dimensional functions. This will be followed up by an explanation of the physics behind the object moving down the slope where factors that are required to calculate the velocity of an object rolling down the slope are identified. Afterwards, some trigonometry and calculus will be applied to derive a function which models the velocity of an object rolling down a complex surface.

RESEARCH

The physics behind an object rolling down a slope

Suppose an object is rolling down a straight slope. The motion is the result of acceleration due to the gravity of the celestial body on the object. The acceleration due to gravity points directly downwards. However, due to the normal force, which is a reaction force to the gravitational pull of the celestial body directed perpendicularly from the contact surface of the object to the slope, the motion of the object is at an angle to the straight surface of the ground as seen in Figure 1, the acceleration of the object is parallel to the gradient of the slope.

We can calculate the acceleration of the object by considering the weight and normal force that acts on the slope. Figure 1 shown below represents the vectors of the normal force, weight, and resultant force of the object down a slope considering the force of friction. The force of friction is not negligible when an object is moving down a slope since friction exists between the surface of the object and the slope. The force of friction is dependent on the materials of the objects which are in contact. However, among the variety of friction types that exist, we will only be considering rolling friction in this system. This is because other types of friction and resisting forces of the motion of the object, such as air resistance, requires mathematical knowledge that exceeds this courses level.

The formula for the force is given by the following equation.

$$F = ma \qquad (1)$$

Where m is the mass of the object and a is the acceleration of the object. However, the force of rolling friction is defined as the following [1]

$$F_f = \mu_r \vec{N} \quad (2)$$

Where μ_r is the coefficient of rolling friction and N is the normal force acted on the object perpendicular to the surface it lies on.

In the diagram below, $\vec{N}$ is the normal force, $\vec{f}$ is the force of rolling friction between the slope and object, $\vec{W}$ is the weight of the object meaning the force of pull of the object from the Celestial body's gravity on the object, and $\vec{A}$ represents the parallel component of the weight of the object down the slope due to the force of gravity. Finally, $\vec{A}$ is the sum of all the forces in this system. We will assume that the object is moving from the beginning so we will consider the constant of rolling friction, μ_r, to calculate the rolling friction for a specific type of material combination.

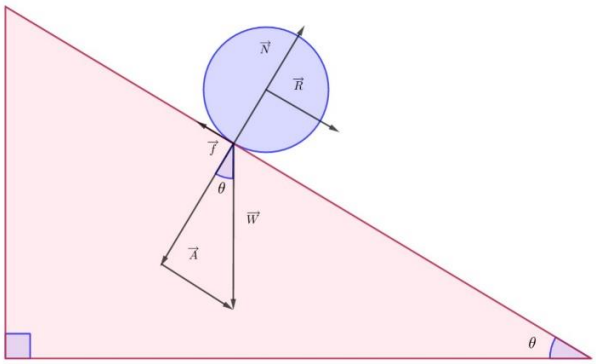


FIGURE 1
THE FORCES ACTING ON A ROLLING OBJECT

As shown in Figure 1, we can see that the weight of the object can be decomposed into two components that are dependent on the angle θ. The component perpendicular to the surface of the triangle which the object lies on cancels out with the normal force of the object $\vec{N}$. $\vec{R}$ is the resultant force of the object, meaning that it is the sum of all forces in this system. Since all of the other forces except the force of friction and $\vec{A}$ caused by the pull of gravity cancel out with other forces in the system, the resultant force will follow the sum of those two forces, thus, the following equation (3).

$$\vec{R} = \sum F = \vec{A} - \vec{f} \quad (3)$$

Figure 1 shows that $\vec{A}$ is the opposite of the angle θ.Thus using basic trigonometry, the force acted on the object due to gravity in the direction parallel to the slope is the following.

$$\vec{A} = mgsin\theta \quad (4)$$

Similarly, using the same method, we will derive an equation for the force of the rolling friction in the case of Figure 1.
From equation (2),

$$F_f = \mu_r \vec{N}$$

From Figure 1,

$$\vec{N} = mgcos\theta$$

Substituting the expression for $\vec{N}$ into the original equation,

$$\vec{f} = \mu_r mgcos\theta \quad (5)$$

Substituting equation (4) and (5) into (3) will give the following expression of the resultant force of the object rolling down a straight slope.

$$\vec{R} = mgsin\theta - mg\mu_r cos\theta \quad (6)$$

Solving for an expression for the parallel component of the weight by setting equation (1) and (6) equal to each other,

$$ma = mgsin\theta - mg\mu_r cos\theta \quad (7)$$

Factoring mg in the RHS and
dividing both sides by m will give the following equation:

$$a = g(sin\theta - \mu_r cos\theta) \quad (8)$$

Equation (8) shows the parallel component of the weight of the object down a straight slope which is tilted at an angle θ to the ground. However, for a complex surface, the angle that exists at a certain point is not constant like it is on a straight slope. Therefore, an expression which takes the change in the angle of the surface at each point into account is required. To find an expression for the change in the angle of the surface, we will first derive an expression for the change in height of the surface to the x, y plane in the direction of the specific vector.

Finding the change in height of a surface to the x, y plane in the direction of a specific vector

Now, consider the complex surface expressed in Figure 2.

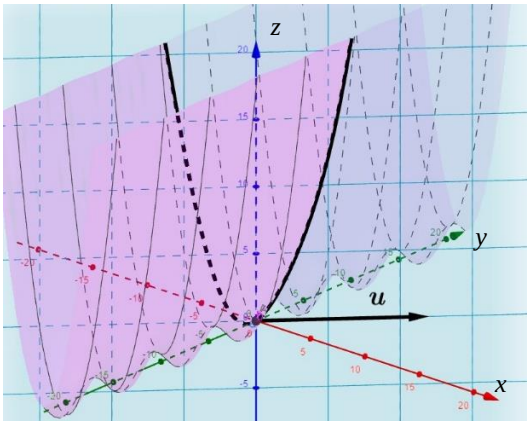


FIGURE 2
CUTTING A COMPLEX SURFACE WITH A PLANE PERPENDICULAR TO THE (x, y) PLANE IN THE DIRECTION OF THE VECTOR

In this case, we will consider the vector $\boldsymbol{u}$ that exists across the x and y plane. This vector is 2-dimensional and its direction does not change in its vertical component in a 3-dimensional space, z. If we are able to cut the surface across the direction of the vector with a plane perpendicular to the x, y plane in a 3-dimensional space, we would be able to find a graph that would represent the change in height from the x and y plane. Since this will be a relationship between 2 variables, the distance across the x, y plane in the direction of

the vector $\boldsymbol{u}$ and the change in height of the slope perpendicular to the x, y plane we know that the function will be 2-dimensional. The diagram on the right shows a plane which exists perpendicularly from the x, y plane intersecting the surface graphed by a random function.

A vector can be thought of as a ratio between the change in value in one direction component x and the other y. Since a vector is straight and not curved, the ratio between the increase in x values and y values are consistent throughout the function. Since the relationship between the two variables is a constant ratio, we can express y in terms of x. Thus, $y = g(x)$. In a 3-dimensional function $f(x, y)$, since y can be expressed in terms of a function of x, the function $f(x, y)$ will now become the following.

$$z = f(x, y) = f\big(x, g(x)\big) = f(x)$$

The function $f(x)$ is a 2-dimensional function that represents the relationship between the value z with the variable x across the line vector $\boldsymbol{u}$. This is true because one input value for x will give a value for y corresponding to that x value. Therefore, for 1 x value, there is only one z value, for all the points across the line $\boldsymbol{u}$, creating a 2-dimensional function of the change in the value of z as the point of focus moves across the vector $\boldsymbol{u}$.

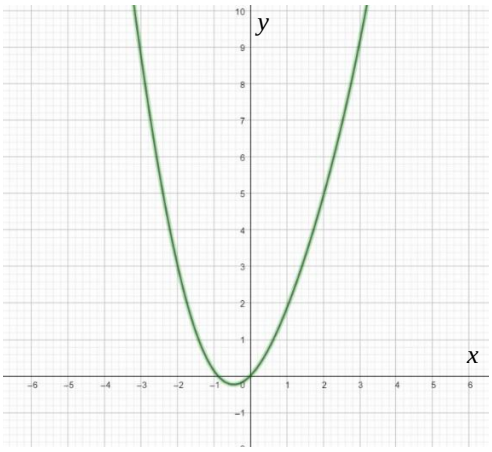


FIGURE 3
THE CHANGE IN HEIGHT OF THE FUNCTION $f(x)$ IN THE DIRECTION OF THE VECTOR $\boldsymbol{u}$

Applying this method to the example interaction shown in Figure 2 will give the intersection shown in Figure 3. This is identical to the intersection of the surface and the plane shown in Figure 3.

Change in the angle at each point of the function

Recall that the angle of the slope with the ground was considered when finding the acceleration of the object down a straight slope. However, in the case of a non-straight surface, the angle at each point on the surface that it creates with a vector parallel to the ground is different. Therefore, this change in angle must be taken into consideration when modelling an accurate change in velocity function.

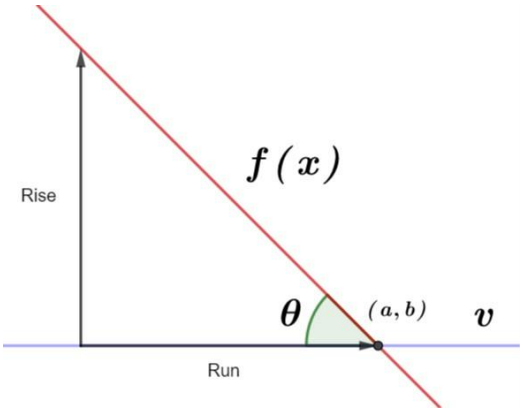


FIGURE 4
$f(x)$ ZOOMED UP TO THE POINT (a, b)

Suppose that we have the function $f(x)$ that represents the change in height of the surface relative to the x, y plane, and suppose we want to find the change in the angle that it creates with $\boldsymbol{v}$ a vector parallel to the x-axis, that intersects at the point (a, b), on the blue line as shown in the figure below. Finding the angle at each point on a changing slope is a necessary method when considering the physics calculations in the next section. This could also be thought as the angle made between a line that is tangent to the function $f(x)$ at the point (a, b), and a line parallel to the x-axis.

Figure 4 shows the function $f(x)$ zoomed up to the point (a, b) on the function. θ is the angle that is created between the function $f(x)$ and the vector $\boldsymbol{v}$ at this very point.

When substituting values into the derivative function, we get a certain ratio between the x values and the y values that show the gradient at that point. Hence, $\frac{dy}{dx}$. However, the gradient is also considered as rise over run in a linear function. Curves can be treated as a line when only a small portion of it is considered. This situation can be in such a form when we zoom into a single point of a function and the curve seems to straighten up to a point where it is indistinguishable from a line, as shown in Figure 4. In this case, the gradient of the function when divided into the rise and run will create a triangle with the function $f(x)$ as shown in Figure 4.

From Figure 4, we can see that the vertical rise and the horizontal run will be opposite and adjacent respectively to the angle θ. By the basics of trigonometric ratios, we know that the $\frac{O}{A}$, the opposite over the adjacent sides of an angle is equal to the tangent of that angle. Hence, we can graph a function that shows the change in the angle θ by taking the inverse of the gradient function $f'(x)$. Hence, the function that represents the change in the magnitude of θ, the angle at the point (a, b) can be expressed in the form of the following function $g(x)$.

$$\theta = g(x) = arctan\,(f'(x)) \qquad (9)$$

Recall equation (8), the acceleration equation that was derived for an object rolling down a straight slope considering the force of rolling friction. However, since the

angle of the slope at a given point on the slope is not constant for a complex surface, we can substitute for the angle with the expression of θ, equation (9), into equation (8).
The acceleration function, equation (8)

$$a = g(sin\theta - \mu_r cos\theta)$$

Substituting for the value of θ with equation (8),

$$a = g(sin[arctan\,(f'(x))] - \mu_r cos[arctan\,(f'(x))]) \quad (10)$$

Thus, equation (10) is now able to show the variation in the acceleration of the object as it slides down the slope.

Manipulating the acceleration function to create a velocity function

Since the acceleration function is the derivative of the velocity function, we know that the velocity function is the integral of the acceleration function. Therefore, integrating equation (10) will give us the velocity function of an object rolling down a complex surface.

$$v = \int g[sin\{arctan\,(f'(x))\} - \mu_r cos\,\{arctan\,(f'(x))\}]\,dx$$
$$= g\int sin\{arctan\,(f'(x))\} - \mu_r cos\,\{arctan\,(f'(x))\}\,dx$$

Simplifying for the value of $sin\{arctan(f'(x))\}$ and $cos\{arctan(f'(x))\}$

Let $arctan\,(f'(x)) = \theta$

$$\therefore sin\{arctan\,(f'(x))\} = sin\,(\theta)$$

The following right-angle triangle can be constructed to represent the value of θ.

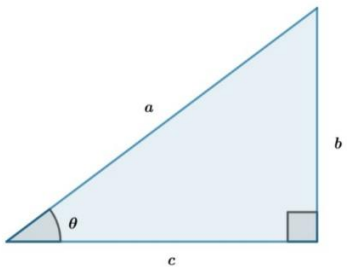


FIGURE 5
A RIGHT-ANGLE TRIANGLE

$$tan\theta = \frac{Opposite}{Adjacent} = \frac{b}{c}$$

$$tan\theta = tan(arctan\,(f'(x))) = f'(x) = \frac{f'(x)}{1}$$

Thus,

$$\frac{b}{c} = \frac{f'(x)}{1}$$
$$\therefore b = f'(x), c = 1$$

By Pythagoras's Theorem,

$$a^2 = b^2 + c^2$$
$$a = \sqrt{b^2 + c^2}$$

Substituting for the values of b and c,

$$a = \sqrt{1 + [f'(x)]^2}$$

Hence,

$$sin\theta = sin\{arctan\,(f'(x))\} = \frac{b}{a} = \frac{f'(x)}{\sqrt{1 + [f'(x)]^2}}$$
$$cos\theta = cos\{arctan\,(f'(x))\} = \frac{c}{a} = \frac{1}{\sqrt{1 + [f'(x)]^2}}$$

Finally, substituting the for sin $(arctan(f'(x)))$ and cos $(arctan(f'(x)))$ in the velocity equation derived previously,

$$v = g\int \frac{f'(x)}{\sqrt{1 + [f'(x)]^2}} - \mu_r \frac{1}{\sqrt{1 + [f'(x)]^2}}\,dx$$

Since the denominators are equal, changing the expression into a single fraction,

$$v = g\int \frac{f'(x) - \mu_r}{\sqrt{1 + [f'(x)]^2}}\,dx \quad (11)$$

Therefore, equation (11) is an equation that is able to calculate the speed of an object rolling down a complex surface in a given direction.

However, equation (11) does not show the true velocity of the object when it is rolling down a slope. This is due to the fact that if this equation is true, a negative slope will give a negative acceleration. This is not the case in the real world since a negative slope will give a positive acceleration since the force of gravity is pulling down on the object. Instead, a positive slope must give a negative acceleration. This can be extrapolated from the fact that hill climbing in cycling is much harder than downhill riding. Therefore, -1 must be multiplied to the equation to change the direction of the acceleration in a manner that if the gradient of a surface is negative, the acceleration must be positive as demonstrated in equation (12).

$$v = -1 \cdot g\int \frac{f'(x) - \mu_r}{\sqrt{1 + [f'(x)]^2}}\,dx$$
$$= -g\int \frac{f'(x) - \mu_r}{\sqrt{1 + [f'(x)]^2}}\,dx$$

However, the direction of the acceleration, the inside of the integral, must be in the opposite direction as well. Thus, placing the negative inside the integral. This does not affect the resultant value of the velocity

$$v = g\int -\frac{f'(x)-\mu_r}{\sqrt{1+[f'(x)]^2}}dx \qquad (12)$$

Recall that we are deriving an equation to find the velocity of the object at a certain point when rolling down a complex surface. Hence, the velocity at a certain point on the surface at a given point would be defined as the sum of the accelerations from where the object began its motion.

Assume that the object will begin its motion at the point where $x = a$ on the vector $\boldsymbol{u}$, and we want to find the velocity of the object when it passes the point $x = b$ on the vector $\boldsymbol{u}$. Therefore, a definite integral can be derived from equation (11) which shows the velocity of an object at a certain point.

$$v = g\int_a^b -\frac{f'(x)-\mu_r}{\sqrt{1+[f'(x)]^2}}dx \qquad (13)$$

This equation is able to give a value for the velocity of the object rolling down a complex surface at a certain point since it is the sum of the accelerations from the beginning of the motion of the object. The velocity of the object is still in the direction of the vector $\boldsymbol{u}$ since there was a substitution for the value of y in the function $f(x,y)$. The change in slope is also considered in this equation from the substitution for the value of θ.

Accuracy check

FIGURE 6
SLOPE USED IN ACCURACY CHECKING EXPERIMENT

To test the accuracy of the formula, the theoretical velocity of an object rolling down a surface will be compared to a real-life value collected from an experiment. A slope was created using a piece of flexible plastic and some tape to keep it in place. The velocity of the object at the moment it reaches the table, thus the maximum velocity of the object will be measured using a motion detector and LabQuest. Figure 6 shows the slope used in the experiment. Since both the object and the slope were made out of plastic, the coefficient of rolling friction was assumed to be 0.4. [2]

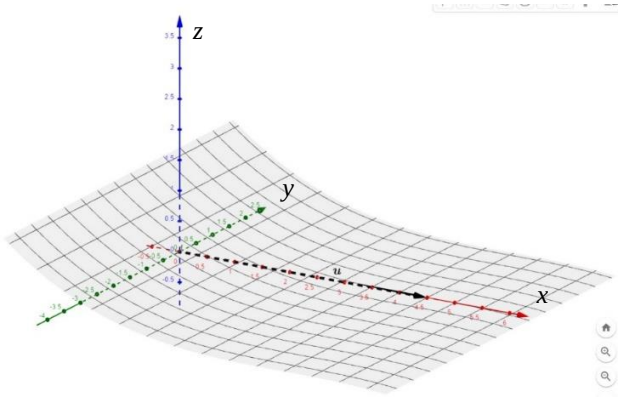


FIGURE 7
THE SLOPE USED IN THE EXPERIMENT MODELLED IN GEOGEBRA

To be able to compare the experimental and theoretical values a function was estimated that will represent the slope in Geogebra. This was done by comparing a side view of the slope while manipulating the function in Geogebra. Figure 7 is a 3D model of the graph created in Geogebra, modelled by the function $f(x,y) = 0.01(-x+4.5)^3$. When calculating the theoretical value using the derived formula, It will be assumed that the object begins its motion at the point $f(0,0)$ and moves across the vector $\boldsymbol{u} = \begin{pmatrix}0\\0\end{pmatrix} + \mu\begin{pmatrix}1\\0\end{pmatrix}$. The velocity of the object will be calculated at the point $f(4.5,0)$.

Solving for the theoretical velocity:

Differentiating $f(x)$

$$\frac{d}{dx}f(x) = -3\cdot 0.01(-x+4.5)^2$$
$$= -0.03(-x+4.5)^2$$

Thus, substituting values into equation (12),

$$v = 9.81\int_0^{4.5} -\frac{-0.03(-x+4.5)^2 - 0.4}{\sqrt{1+[-0.03(-x+4.5)^2]^2}}dx$$

Using GeoGebra to solve for the area under the curve, a to find a value for the finite integral in the equation above,

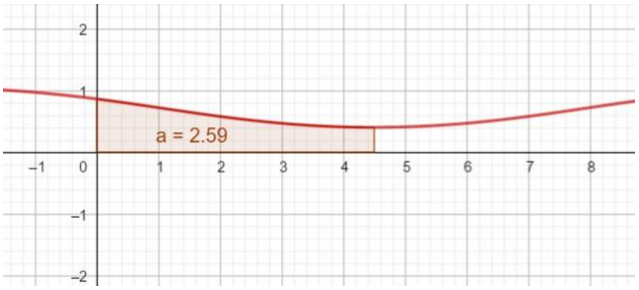


FIGURE 8
SOLVING FOR THE AREA UNDER THE CURVE USING GEOGEBRA

From the area under the curve, a shown in Figure 8,

$$\therefore v = 9.81\cdot 2.59$$
$$= 25.4079\ units\ s^{-1}$$

Assigning units to this theoretical velocity to enable comparison with the measured value. Since the horizontal

length of the system was 50cm, multiplying the theoretical velocity by $\frac{50}{4.5}$, the ratio between the measured length of the system and the units on the cartesian plane, to scale the 3D simulation up to the dimensions of the slope created.

$$\begin{aligned} v &= 25.4079 \cdot \frac{50}{4.5} \\ &= 282.31 cms^{-1} \\ &= 2.8231 ms^{-1} \end{aligned}$$

The measurement collected from the ultrasound detector was $2.695 ms^{-1}$. Although the theoretical value is not equal to the measured value, numerous uncertainties existed in the experiment since air resistance was considered negligible and the coefficient of rolling friction was not measured but estimated. Since the measured values were similar to a great extent, it can be concluded that the formula is accurate to a certain extent.

CONCLUSION

Beginning with physics behind an object rolling down a straight slope, an equation which takes the change in gradient of a complex surface was derived. The formula gives a theoretical value of the velocity of a rolling object.

From the accuracy check performed by an experiment shows that the formula is capable of estimating the velocity of a rolling object down a complex surface in real life with a certain amount of accuracy. Thus, there is a possibility that this equation will be able to output values as accurate as a computer simulation.

However, this equation has its limitations since it is only applicable under certain assumptions of the physical world and cannot be applied in certain cases. This equation does not take into account the law of conservation of energy. Therefore, depending on where the object begins its motion and its initial velocity, this equation will give false values for the speed of the ball. For example, if the object is placed where the height of the surface relative to the x, y plane is a minimum, (Gilliland and Dupree) and its kinetic energy is less than the work required to travel up the surface to the next maximum point, the velocity that this equation gives cannot be trusted since it will defy the law of conservation of energy.

Another assumption that is made is that the object is always in contact with the material when travelling across a specific path of a surface. If the kinetic energy of the object is greater than the gravitation potential energy that it will gain at the maximum point of the surface, the object is likely to take off on a projectile which this equation cannot take into account.

Furthermore, this equation is applicable under the assumption that the gravitational acceleration is perpendicular to the x, y – plane.

Otherwise, the equation would be able to simulate the theoretical velocity of a bicycle rolling down hill, which was the main aim of developing this equation. Using this equation, one would be able to analyse a slope and determine at approximately which point the breaks should be applied so the bike doesn't go too fast and leads to the rider losing control of the bike.

Although the aim was to model the velocity of a bike moving down a slope, this equation can be applied in some other situations such as space exploration and in automated robots. This equation is able to give a theoretical value of the velocity of an object moving across a complex surface. Since this equation can simulate the velocity of an object going down a complex surface, it will mean that it would be able to calculate the velocity of a machine such as a road vehicle or a robot with tires that is going down a complex surface. Since there are limits to which the brakes can be applied safely to the rotating wheels, by simulating the velocity of the robot going down the surface, we can determine whether it is safe or not for the path to be taken. Technologies such as doppler be able to create a map of the surface, and in combination with this equation, it will allow the AI or person in control of the instrument, whether to go down the path or not. [3] Therefore, this could prevent any possible damages to the expensive machines such as robots that move around surfaces on another planet.

REFERENCES

[1] Engineering ToolBox, (2008). *Rolling Resistance*. [online] Available at: https://www.engineeringtoolbox.com/rolling-friction-resistance-d_1303.html, Accessed October 1st.

[2] Coefficient of Friction, Rolling Resistance, Air Resistance, Aerodynamics. [online] Available at: http://www.tribology-abc.com/abc/cof.htm. Accessed 19 Apr. 2019.

[3] Gilliland, Ronald L., and A. K. Dupree. 'First Image of the Surface of a Star with the [ITAL]Hubble Space Telescope[/ITAL]'. The Astrophysical Journal, vol. 463, no. 1, May 1996, pp. L29–L32. Institute of Physics, doi:10.1086/310043.

An Investigation Into the Sodium-metal Battery As an Alternative For Globally High Demand Lithium-ion Batteries

Mayuko Yamamoto
Ritsumeikan Uji Senior High school, Uji City, Country
Email: mayuko.yamamoto2020@gmail.com

***Abstract* – Sodium-ion batteries have attracted attention by researchers as a more suitable alternative to lithium-ion batteries. However, due to its larger ionic radius compared to lithium-ions, a suitable anode that does not hinder the intercalation process is yet to be found. In this work, I demonstrate that issue of large ionic radius of sodium can be avoided with sodium-metal battery. Sodium metal battery involves sodium metal and copper wire, which create electric field. The experimental observations show that the electric field can be easily strengthened by increasing the concentration of sodium chloride solution in which two metals are put (indirectly for sodium metal as it reacts with water). The maximum p.d. obtained by the sodium-metal battery is comparable to that produced by current lithium-ion battery, further supporting the possible suitability of sodium-metal battery as an alternative for lithium-ion battery. This allows to overcome the current global and local concern of lithium-scarcity, encouraging to shift from fossil fuels to renewable energies and help develop technologies.**

Key Words – Lithium-ion batteries, Sodium-ion batteries, Sodium-metal batteries, potential difference, Sodium chloride solution

INTRODUCTION

Lithium-ion batteries are present in many devices, including phones, cars, computers, tablets, and hearing aids [1]. Furthermore, the recent focus on producing electric vehicles and portable electronics is diversifying the use of rechargeable batteries[2]. Moreover, as the world shifts away from fossil fuels, rechargeable batteries become a key to store the energy produced from renewable sources[3]. However, in contrast to the growing needs of lithium-ion batteries, the demand of lithium is expected to surpass the supply after 2024, unless "more projects are commissioned" to produce more lithium[4]. In Japan, 100% of lithium carbonate and lithium hydroxide used in lithium-ion batteries are imported from overseas[5]. Hence, the electrical industry of Japan would be particularly and largely affected if the price of lithium rises due to the increasing demand.

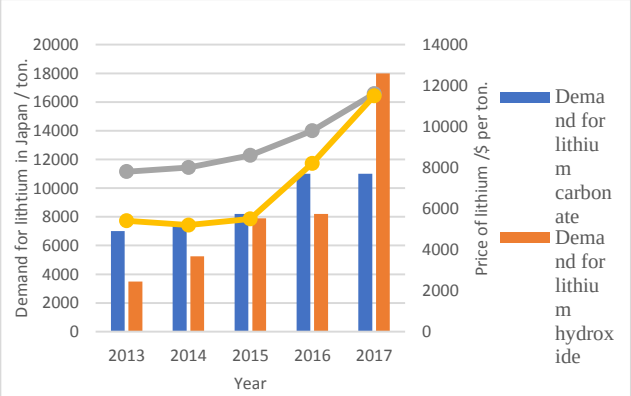


Figure 1: Demand and price of import for lithium (lithium carbonate and lithium hydroxide) in Japan created from data taken from Mineral Resources Material Flow

As figure 1 shows, the price of lithium has increased with a rapidly increasing demand for lithium. Without finding an alternative for lithium-ion batteries, Japan will not be able to remain competitive in the electrical industry.

Therefore, amid the concerns regarding the scarcity of lithium, I decided to conduct an investigation on possible alternatives to lithium-ion batteries. (The more defined research question of this paper will be stated after all the background researches are explained.)

BACKGROUND INFORMATION

I. Lithium-ion batteries

Since lithium-ion battery has entered commercial market in 1991[6], it has been used as the most popular rechargeable battery due to advantages, such as high energy density, the property of not having memory effect, and its capability of undergoing hundreds of recharging cycles[7]. Although it also has weaknesses such as degradation and its sensitiveness to

[1] "The Batteries Report 2018"
[2] ibid
[3] Qi Li, 2017
[4] "End in Sight to Near-Term Lithium Supply Shortages", 2017
[5] 鉱物資源マテリアルフロー2018 [Mineral Resources Material Flow 2018].
[6] "Lithium-Based Batteries Information.", 2018
[7] Marshall, Brain, 2016

high temperatures, these are considered insignificant compared to the large advantages that it has[8].

Figure 2 shows the composition of one of the most common lithium-ion batteries, cobalt oxide battery, which has lithium cobalt oxide as the cathode and graphite as the anode [9]. The conductive metal layers, aluminium and copper function as collectors to aid the electrodes in collecting and distributing electrons[10]. An electrolyte is a transport medium which is permeable for lithium ions but not for electrons[11]. Through this electrolyte, lithium ions move between cathode and anode so that it can neutralise the charge built up by electrons (which travel through the external circuit) to prevent electrons from repelling each other[12]. Furthermore, the separator, which is also permeable for ions, is placed between cathode and anode, to prevent short circuit when electrolyte evaporates due to high temperatures in abnormal conditions[13].

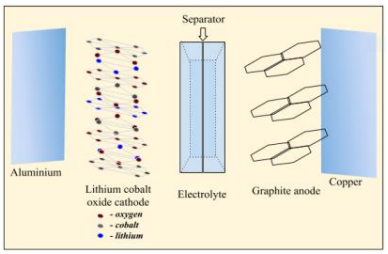


Figure 2: Simplified diagram of a lithium-ion battery made by me on Google doc

I. Sodium-ion battery, and its limitations

Since 2012, sodium-ion battery has been attracting attention from scientists globally as a possible alternative for lithium-ion battery, as seen from the increasing number of publications on it[14]. This is because sodium has similar chemical properties as lithium as it is an element one period below on the periodic table. Moreover, the abundance of sodium is as high as 2.36×10^4 mg/L, whereas that of lithium is merely 20 mg/L[15]. This also means that the cost of the raw material, sodium, for this type of battery is cheap, too. In fact, the cost of mining sodium is approximately $150 US dollars per tonne, while mining lithium costs $150,000 per tonne[16].

However, one of the biggest problems with sodium-ion batteries is the difficulty with finding a suitable anode. Although graphite is commonly used as an anode in lithium-ion battery as shown in figure 2, a graphite anode is not functional in sodium-ion battery due to the larger ionic radius of sodium (116pm) compared to lithium (90pm)[17], which hinders the intercalation process[18]. Therefore, many researches have focused on anode material such as research conducted by Komaba et al. on hard-graphite[19], Jache and Adelhem on expanded graphite[20], and others[21,22]. However, a suitable anode is still at the stage of development and the commercial application of sodium-ion battery is yet to be implemented.

Hence, the focus of investigation must be broadened out of sodium-ion battery with conventional graphite anode.

II. Sodium-metal battery

Sodium-metal battery is a battery type that has similar structure with sodium-ion batteries, except it does not involve a graphite-anode. 'Sodium metal battery' means that sodium metal itself is an electrode, instead of it being a compound. This allows to avoid the difficulty with finding a suitable material for battery anode in sodium-ion battery. Furthermore, this sodium metal battery also holds the advantage of using a sodium which is significantly more abundant than lithium.

However, despite this advantage of sodium-metal batteries and successful results obtained from researches done by Cohn, Adam P., et al.[23], researches into sodium-metal batteries are very little compared to sodium-ion batteries.

Therefore, I decided to focus my research on sodium-metal battery, particularly, on one possible factor that might affect the potential difference that can be produced from a battery with the research question: **"What is the effect of concentration of sodium chloride solution on the potential difference (p.d.) produced across the sodium-metal battery model?"**

EXPERIMENTAL SET-UP

Figure 3 shows a diagram of the battery model. This was adopted from a video posted on YouTube by Ian Berry[24].The video used a very basic concept of sodium battery model. Therefore, I modified the model by conducting multiple preliminary experiments to create the best method for obtaining relevant and reliable data for this research inquiry.

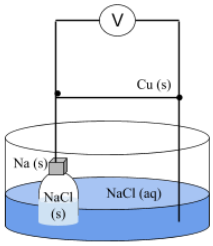


Figure 3: experimental setup drawn on Google

I. Justification of model and experimental design

Rechargeable batteries are similar to capacitors, and the idea of capacitors can help understand the mechanism

[8] "Lithium-Based Batteries Information.", 2018
[9] Kazda, Tomáš, and Petr Vanýsek. 2016.
[10] Tablante, Teddy, 4 Mar. 2019
[11] ibid
[12] Ibid
[13] ibid
[14] Sawicki, Monica, and Leon L. Shaw, 2015
[15] Todd Helmenstine, 2018
[16] Brooks Hays, 2017
[17] "Periodic Trends"
[18] Yasuyuki Kondo, et al., 2019
[19] Shinichi Komaba, et al., 2011
[20] Birte Jache and Philipp Adelhelm, 2014
[21] Xinwei Dou, et al., 2019
[22] Hongshuai Hou, et al., 2017
[23] Cohn, Adam P., et al., 2017
[24] Ian Berry, 2016

of my battery model, which has two conducting metals, copper and sodium, and two dielectrics, NaCl crystal and NaCl solution.

Firstly, electric potential difference (p.d.) is the difference in the electric potential energy between two points[25]. This p.d. then causes current through a circuit as the electrons are repelled from the negative side and attracted to the positive side and flows through the complete circuit.

Furthermore, p.d., is work done when a charge moves through an electric field[26], which are related by equation:

Equation 1[27]: $V = Ed$, where E is the electric field strength, V is the p.d., and d is the separation between the conductors.

This equation indicates that p.d. is proportional to electric field strength and as the electric field strength increases, the p.d. increases.

Finally, electric field strength is dependent on the relative electric permittivity of the dielectric between the conductors.

The relative electric permittivity of a material is its ability to polarise to the external electric field[28] which are related by an equation:

Equation 2[29]: $E = \frac{D}{\varepsilon_0 \varepsilon_r}$, where D is the flux density, ε_0 is the absolute permittivity of vacuum which is constant, and ε_r is the relative permittivity of the medium.

This suggests that the electric field strength (and hence p.d.) across the metals is inversely proportional to the relative permittivity of the dielectrics between them.

Given this idea, I will now move onto explaining how each of the component of the battery model contributes to the p.d. across the metals.

II. P.d. and Electrodes

Batteries undergo electrochemical processes which involves the oxidation of the battery anode due to the removal of electrons and reduction of the battery cathode due to the gain in electrons. The build-up of charge on electrodes then creates p.d. and causes the electrons to move from the negatively charged end to the positively charged end, creating current.

In my battery model, the anode is the sodium metal and the cathode is the copper wire. This can be understood from the electrode potentials of sodium and copper: sodium has a value of -2.71V and copper has a value of +0.15V[30]. This suggests that sodium is more likely to be oxidised than copper, thus becoming an anode, making copper wire a cathode.

For this component, the areas of metal which are in contact with the dielectric ultimately contribute to the p.d. obtained. This relationship can be understood with the equation:

Equation 3[31]: $D = \frac{Q}{A}$, where Q is the charge built-up and A is the area of the conductors in contact with the dielectric.

Equation 1 and 2 have shown the proportionality between the flux density and electric field strength, and the proportionality between electric field strength and p.d. respectively. Therefore, these equations together show how the areas of conductors in contact with dielectrics and the p.d. are related. Hence, the areas of sodium metal and copper was controlled, which is further explained in the experimental design section.

III. Use of NaCl crystal as the prevention as the prevention of Coulomb explosion

Next, the battery model includes two dielectrics, one of which is NaCl crystal. The main function of NaCl crystal in the model is to elevate the sodium metal to prevent "coulomb explosion" from happening without disturbing the p.d.

Firstly, when an alkali metal and water are in direct contact, coulomb explosion takes place, which is an "explosive alkali metal-water reaction that leads to rapid migration of electrons from the metal into water"[32]. When I tried to measure p.d. while sodium metal and NaCl solution are in direct contact, coulomb explosion occurred extremely rapidly that p.d. was unable to be detected. In addition, since the reaction is explosive, a battery model in which alkali metal and water are in direct contact would be practically useless.

Therefore, I tried to prevent direct contact of the water and the metal by elevating the metal above NaCl solution with NaCl crystal.

Again, NaCl crystal is a dielectric with a dielectric constant of 6.1[33], and an electric field can travel through it[34]. Therefore, this allows p.d. to be measured without coulomb explosion occurring.

NaCl crystal is used because even if it dissolves in the solution during the experiment, it will not affect the composition of the solution, which helps reducing random errors.

IV. NaCl solution and p.d. obtained

[25] "Electric Potential Difference"
[26] "Electric Potential Difference"
[27] "Permittivity and Relative Permittivity or Dielectric Constant." 2019
[28] Nave, Carl Rod.
[29] "Permittivity and Relative Permittivity or Dielectric Constant." 2019
[30] Haynes, Willian M., et al., editors., 2014
[31] "Permittivity and Relative Permittivity or Dielectric Constant." 2019
[32] Phillip Mason et al., 2015
[33] "Dielectric Constant Values."
[34] Richard Williams, 1963

The final component is the second dielectric, NaCl solution. This largely accounts for the p.d. obtained and it acts like an electrolyte in the model.

Firstly, water is a dielectric with a relative permittivity of 80.2 at 58° F[35] (20 °C). It has been introduced earlier that permittivity is a measure of the material's ability to polarize to an external electric field. This polarization actually creates a new electric field opposing the original external field, and results in the reduction of the field strength between the conductors [36]. An example of the effect of polarization is shown with water molecules in figure 4.

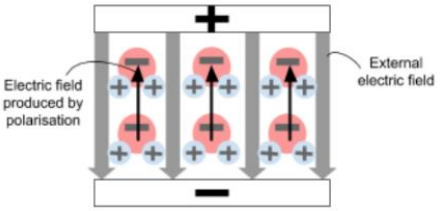


Figure 4: the effect of polarisation with water molecules created by me on Google doc

Additionally, part of the research by D H Dagani et al. shows that the permittivity of the water decreases as NaCl concentration increases [37]. Further research by Nir Gavish and Keith Promislow shows that NaCl ions in the solution creates local electric field and causes polar water molecules to align with this field creating a "hydrogen shell" [38] as shown in figure 5.

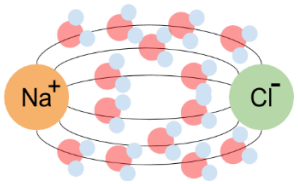


Figure 5: water molecules aligned to the electric field created on Google Doc by me

Moreover, they found that the permittivity of NaCl solution decreases with concentration by the process called "dielectric decrement", which is expected to be linear for dilute solutions ($\leq 1.5M$) and saturate at higher concentrations[39]. This is because in concentrated solutions, larger number of NaCl ions polarize the water molecules to their field, reducing water molecules' ability to polarize (permittivity) to the external electric field. This ultimately means that increasing the NaCl concentration results in the reduction of the field strength opposing the external field, and hence increasing the external electric field strength (hence p.d.).

HYPOTHESIS

Overall, the research shows how the dielectric decrement is linear at lower concentrations and saturates at higher concentrations, the electric field strength is inversely proportional to the permittivity, and electric field is directly proportional to the p.d.

Therefore, I hypothesise that p.d. produced across the battery increases rapidly at lower concentrations and stabilises at higher concentrations. However, the extent to which this occurs will be under investigation throughout the experiment.

EXPERIMENT

Materials:

- Sodium metal
- NaCl crystal
- NaCl solutions (0M, 1M, 2M, 3M, 4M)

Apparatus

- Data logger with p.d. sensor
- Copper wire
- Crocodile clip

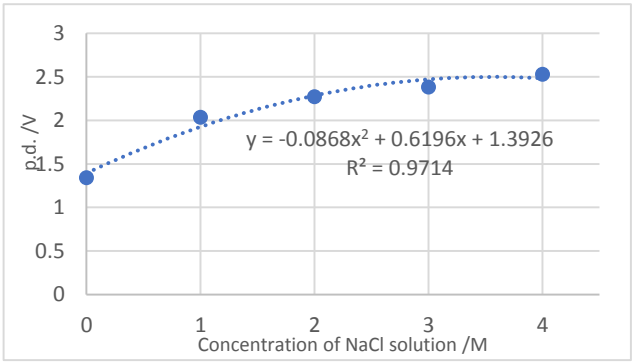
RESULTS

I. Raw data table

Conc. of NaCl solution /M	Avg. p.d. measured within an interval of 5sec./V					
	1	**2**	**3**	**4**	**5**	Avg
0.00	1.441	1.078	1.483	1.570	1.095	1.333
1.00	2.016	2.084	2.004	1.959	2.109	2.034
2.00	2.261	2.263	2.266	2.290	2.269	2.270
3.00	2.437	2.403	2.359	2.353	2.355	2.381
4.00	2.543	2.470	2.509	2.556	2.571	2.530

Table 1: raw data table presenting the p.d. measured on LoggerPro within the duration of 5 seconds

II. Processed data

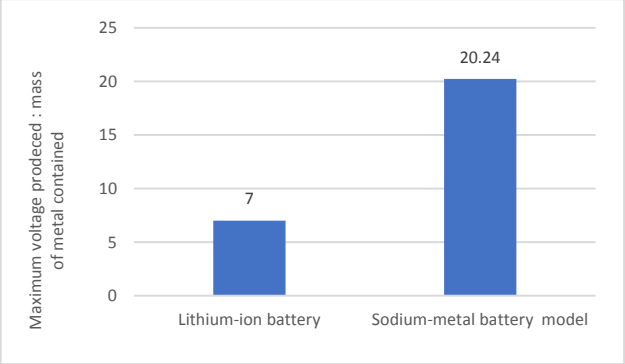


Graph 1: Different concentrations of NaCl solution plotted against the average p.d. measured

[35] "Dielectric Constant Values."
[36] Nave, Carl Rod.
[37] Gadani, D H, et al. 2012
[38] Gavish, Nir, and Keith Promislow, 2016
[39] Gavish, Nir, and Keith Promislow, 2016

Graph 1 shows an overall increase in p.d. as the concentration of NaCl increases as hypothesised. Since p.d. increased greatly at lower concentrations but it seemed to maximise at about 4M, a quadratic trend line, $y = 0.0868x^2 + 0.619x + 1.3926$, is plotted, which has a high R^2 value of 0.9714. This equation also suggests that the rate of change in p.d. decreases linearly with concentration. This can be understood from the research that trend of the dielectric decrement is rapid at lower concentrations and saturates at higher concentrations. The processes of dielectric decrement and hydrogen shell resulted in decrease of permittivity with increase in NaCl concentration which led to increasing external electric field strength and hence increasing the p.d. across the model.

In this battery model, average maximum p.d. obtained was 2.530V with 4M NaCl solution. Graph 2 compares this maximum p.d. obtained from approximately 0.125g of sodium metal and 18650 lithium-ion battery cell which has approximately 0.6g of lithium content per cell[40] and produces a maximum p.d. of 4.2V per cell[41]. The bar represents the ratio of the maximum voltage produced to the mass of metals (sodium and lithium) for sodium-metal battery model and lithium-ion battery.



Graph 2: The comparison between the ratios of maximum voltage produced to metal contained, between 18650 lithium-ion battery and sodium-metal battery model

From graph 2, it is evident how the maximum p.d. produced per mass of sodium-metal battery model is nearly three times that of lithium-ion battery cell. This clearly visualises advantage that sodium-metal battery has over lithium-ion battery in terms of the maximum p.d. produced. This initial result strongly supports the potential of sodium-metal battery as an alternative to lithium-ion battery

CONCLUSION

Many of the electric devices are relying on lithium-ion batteries in the current world, and the recent focus on producing electric vehicles and finding a way to store renewable energies has further enhanced the demand of rechargeable batteries. However, the foreseen scarcity of lithium and price that rises corresponding to its demand are causing concerns in the world, especially in Japan where lithium is imported and hence largely affected by the price increase. The sodium-ion battery, which has been suggested as a possible alternative for lithium-ion batteries has a major problem that its larger ionic radius hinders the intercalation process to the graphite anode (which is a common anode used in lithium-ion batteries).

This paper aimed to offer insights into the effectiveness of sodium-metal battery as an alternative to lithium-ion battery through experiment with a sodium-metal model. The result indicates that there is potentially greater p.d. produced using sodium metal in a battery than lithium per unit mass of each metal.

However, much more experimentation is needed to validate this finding as discussed following an analysis of the data. Furthermore, more research into the alternative rechargeable batteries to lithium-ion batteries must be conducted to meet the current global power demands and to aid the use of renewable energy sources and reduce our reliance on fossil fuels.

REFERENCES

[1,2] "The Batteries Report 2018." The Advanced Rechargeable & Lithium Batteries Association, Recharge, www.rechargebatteries.org/wp-content/uploads/2018/05/RECHARGE-The-Batteries-Report-2018-April-18.pdf.

[3]Li, Qi, et al. Solar Energy Storage in the Rechargeable Batteries. 2017, Solar Energy Storage in the Rechargeable Batteries.

[4] "End in Sight to Near-Term Lithium Supply Shortages." BloombergNEF, 17 Oct. 2017, about.bnef.com/blog/end-sight-near-term-lithium-supply-shortages/.

[5] 鉱物資源マテリアルフロー2018 [Mineral Resources Material Flow 2018]. Japan Oil, Gas and Metals National Corporation, 2019, pp. 129–143, mric.jogmec.go.jp/wp-content/uploads/2019/03/material_flow2018_Li.pdf

[6,8] "Lithium-Based Batteries Information." Lithium-Based Batteries Information – Battery University, 1 June 2018, batteryuniversity.com/index.php/learn/article/lithium_based_batteries.

[7] Brain, Marshall. "How Lithium-Ion Batteries Work." HowStuffWorks, HowStuffWorks, 14 Nov. 2006, electronics.howstuffworks.com/everyday-tech/lithium-ion-battery.htm.

[9] Kazda, Tomáš, and Petr Vaný sek. Lithium Batteries as Electrochemical Sources of Energy. The

[40] "Is Lithium-Ion the Ideal Battery?"
[41] Sears, George, 2017

Electrochemical Society Interface, 2016, Lithium Batteries as Electrochemical Sources of Energy.

[10,11,12,13] Tablante, Teddy, director. The Life of a Smartphone Battery. YouTube, Branch Education, 4 Mar. 2019, www.youtube.com/watch?v=G5McJw4KkG8

[14] Sawicki, Monica, and Leon L. Shaw. Advances and Challenges of Sodium Ion Batteries as Post Lithium Ion Batteries. Royal Society of Chemistry, 2015, Advances and Challenges of Sodium Ion Batteries as Post Lithium Ion Batteries.

[15] Helmenstine, Todd. "Abundance of Elements in Earth's Crust - Periodic Table and List." Science Notes and Projects, 23 May 2018, sciencenotes.org/abundance-of-elements-in-earths-crust-periodic-table-and-list/.

[16] Hays, Brooks. "Sodium Could Replace Lithium for More Cost-Efficient Battery Storage, Scientists Say." UPI, 10 Oct. 2017, 11:30, www.upi.com/Science_News/2017/10/10/Sodium-could-replace-lithium-for-more-cost-efficient-battery-storage-scientists-say/7681507639505/.

[17] Boundless. "Periodic Trends." Lumen, Lumen, courses.lumenlearning.com/trident-boundless-chemistry/chapter/periodic-trends/.

[18] Kondo, Yasuyuki, et al. Investigation of Electrochemical Sodium-Ion Intercalation Behavior into Graphite-Based Electrodes. 2019, Investigation of Electrochemical Sodium-Ion Intercalation Behavior into Graphite-Based Electrodes.

[19] Komaba, Shinichi, et al. Electrochemical Na Insertion and Solid Electrolyte Interphase for Hard‐Carbon Electrodes and Application to Na‐Ion Batteries. 2011, Electrochemical Na Insertion and Solid Electrolyte Interphase for Hard‐Carbon Electrodes and Application to Na‐Ion Batteries.

[20] Jache, Birte, and Phillip Adelhelm. Use of Graphite as a Highly Reversible Electrode with Superior Cycle Life for Sodium‐Ion Batteries by Making Use of Co‐Intercalation Phenomena. 2014, Use of Graphite as a Highly Reversible Electrode with Superior Cycle Life for Sodium‐Ion Batteries by Making Use of Co‐Intercalation Phenomena.

[21] Dou, Xinwei, et al. Hard Carbons for Sodium-Ion Batteries: Structure, Analysis, Sustainability, and Electrochemistry. 2019, Hard Carbons for Sodium-Ion Batteries: Structure, Analysis, Sustainability, and Electrochemistry.

[22] Hou, Hongshuai, et al. Carbon Anode Materials for Advanced Sodium-Ion Batteries. 2017, Carbon Anode Materials for Advanced Sodium-Ion Batteries.

[23] Cohn, Adam P., et al. Anode-Free Sodium Battery through in Situ Plating of Sodium Metal. American Chemical Society, 2017, Anode-Free Sodium Battery through in Situ Plating of Sodium Metal, pubs.acs.org/doe/abs/10.1021/acs.nanolett.6b05174

[24] Berry, Ian, director. Sodium Battery. YouTube, 18 Apr. 2016, www.youtube.com/watch?v=Ks4nE_EdkD0&t=203s.

[25,26] Electric Potential Difference." The Physics Classroom, www.physicsclassroom.com/class/circuits/Lesson-1/Electric-Potential-Difference.

[27,29,31] "Permittivity and Relative Permittivity or Dielectric Constant." Electrical4U, 2 Jan. 2019, www.electrical4u.com/permittivity-and-relative-permittivity-or-dielectric-constant/.

[28,36] Nave, Carl Rod. "Polarization of Dielectric." HyperPhysics, hyperphysics.phy-astr.gsu.edu/hbase/electric/dielec.html.

[30] Haynes, Willian M., et al., editors. CRC Handbook of Chemistry and Physics. 94th ed., Taylor & Francis Group, LLC, 2014, CRC Handbook of Chemistry and Physics, books.google.co.jp/books?id=bNDMBQAAQBAJ&printsec=frontcover&source=gbs_ViewAPI&redir_esc=y#v=onepage&q&f=false.

[32] Mason, Phillip E., et al. Coulomb Explosion during the Early Stages of the Reaction of Alkali Metals with Water. Macmillan Publishers Limited., 2015, Coulomb Explosion during the Early Stages of the Reaction of Alkali Metals with Water, www.nature.com/articles/nchem.2161.

[33,35] "Dielectric Constant Values." Clipper Controls, www.clippercontrols.com/pages/Dielectric-Constant-Values.html#V.

[34] Williams, Richard. High Electric Fields in Sodium Chloride. Pergamon Press, 1963.

[37] Gadani, D H, et al. Effect of Salinity on the Dielectric Properties of Water. vol. 50, Indian Journal of Pure & Applied Physics, 2012

[38,39] Gavish, Nir, and Keith Promislow. Dependence of the Dielectric Constant of Electrolyte Solutions on Ionic Concentration - a Microfield Approach. 2016, Dependence of the Dielectric Constant of Electrolyte Solutions on Ionic Concentration - a Microfield Approach.

[40] "Is Lithium-Ion the Ideal Battery?" Battery University, batteryuniversity.com/index.php/learn/archive/is_lithium_ion_the_ideal_battery.

[41] Sears, George. "An Introduction to Lithium 18650 Batteries." ElectricBike.com, 3 May 2017, www.electricbike.com/introduction-lithium-18650-batteries/.

An Analysis On The Effect Of Changing Propeller Diameter And Pitch On The Propulsive Efficiency

Sosuke Kojima
Ritsumeikan Uji Junior and Senior High School, Uji, Japan
Email: sosuke.kojima.ritsu@gmail.com

***Abstract* – This research answers the question "What is the effect of changing the propeller diameter and pitch of a GF model propeller on its propulsive efficiency under static thrust?". To measure the efficiency, each propeller with diameters ranging from 6 to 10 inches as well as pitch ranging from 3 to 7 inches were connected to a 12-volt motor, measuring its thrust when 5 volts were applied. The propulsive efficiency was then calculated using collected values of thrust, torque, as well as the calculated value of the propeller disc area. The data collected was analyzed using Excel's graphing tool, alongside accepted literature values. Through the analysis, it was concluded that increasing the propeller diameter and pitch increases its propulsive efficiency in a logarithmic way, however further research into this topic is required to test larger ranges of propeller pitch and diameter.**

Key Words – Physics, Aerodynamics, Propulsive Efficiency, Static Thrust

INTRODUCTION

The reason for researching about the topic of efficient propeller designs comes from the recent expansion in drone technology, and the expansive future uses of drones. Drones are used in areas such as photography, data collection, transportation and even rescue operations, however one major flaw is the run-time of each drone. On average, commercial drones can only fly for 25 minutes at their maximum speed [1], which is very short if practical use is concerned. Therefore, in this research, I will be focusing on the effects of changing the pitch and diameter of a propeller on the thrust created and using this value to ultimately calculate the efficiency of the propeller.

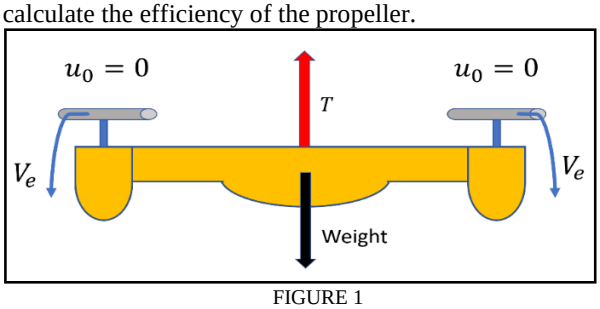


FIGURE 1
FREE BODY DIAGRAM OF DRONE DURING HOVER

In this experiment, the propellers will be operating at static thrust, meaning there is no free-flowing velocity in front of the propeller. This situation is the same condition as a drone hovering, as the drone is stationary, with no free-flowing wind moving towards the drone.

The operation of drones in static thrust is depicted in Figure 1, where T refers to the thrust produced, u_0 the free-flow velocity, and v_e the exit velocity of air. Finding an efficient propeller under static thrust is useful for future research into drone propellers, as drones often operate in hovering conditions when flying. There are also generally fewer research done on performance of static thrust and small-scale propellers, therefore this experiment was conducted. All propellers used in this experiment were manufactured by "Master Airscrew", using their GF models.

BACKGROUND RESEARCH

I. Calculating Propulsive Efficiency

Unlike dynamic thrust where the efficiency can be calculated using values of the inflow velocity of air, static thrust has a more complicated equation to calculate the propulsive efficiency, which is defined as [2]:

$$\eta_{Prop} = \eta_{Propeller} \times \eta_{Motor} \quad (1)$$

The efficiency of the propeller is defined as the ratio between the power induced and the required mechanical power. The power induced is thrust multiplied by the velocity of the outflow air, which by the definition given by Theys Bart and Aron Brezina becomes [3][4]:

$$P_{induced} = T^{1.5} / \sqrt{2\rho A} \quad (2)$$

Where A is the propeller disk area and ρ the air density. The required mechanical power is torque (τ) multiplied by angular velocity (ω), therefore:

$$\eta_{Propeller} = T^{1.5} / \ \tau \times \omega \times \sqrt{2\rho A} \quad (3)$$

The efficiency of the motor is defined by the ratio of mechanical power and inputted power.

$$\eta_{motor} = \tau \times \omega \ / \ P_{in} \quad (4)$$

Hence combining (3) and (4), the overall propulsive efficiency is defined as:

$$\eta_{Prop} = T^{1.5} / \ V \times I \times \sqrt{2\rho A} \quad (5)$$

Air density was calculated to be 1.204 kgm^{-3} using the equation of state given by NASA as [5]:

$$\rho = p/RT_k \quad (6)$$

Where p is pressure, T_k temperature in kelvin and R gas constant (for air 0.286).

II. Thrust and Drag

Now knowing that the change within the variables in (5) will affect efficiency, thrust was chosen as the independent variable, as it is directly related to how the propellers are designed. To decide what to change in order to increase the thrust produced by the propellers, the following equation describing thrust was examined [6]:

$$T = \frac{1}{2} * \rho * A * [V_e^2 - V_0^2] \quad (7)$$

Where V_e is the exit velocity of air, and V_0 the velocity of the aircraft. Along with thrust, drag will also be induced on the propeller that counteracts the thrust produced. Therefore, researching about the drag force, the drag force can be modelled by the following equation [7]:

$$F_D = \frac{1}{2} C \rho A v_0^2 \quad (8)$$

Where A is the area of the object facing the motion of the fluid, and C the drag coefficient.

III. Propeller Diameter

Equation (7) suggests that increasing the propeller disk area would increase the thrust produced, and in order to increase the propeller disk area, the diameter of the propeller could be increased. The propeller diameter is defined as "the diameter of the circle in which the propeller rotates" [8], therefore increasing this diameter would result in a greater disk area.

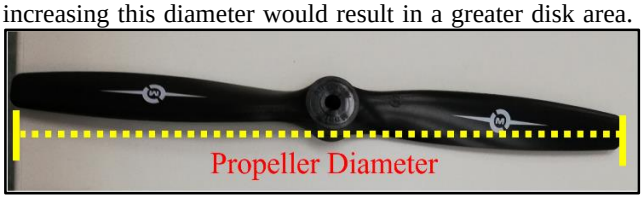


FIGURE 2
DESCRIPTION OF PROPELLER DIAMETER

However, considering (8), it is clear that although the increase in propeller diameter increases the thrust produced, it also increases the area of the object facing the liquid, therefore increasing drag at the same time. This increase in drag is exponential for objects that are in motion, and so the drag force experienced by the propeller increases exponentially, thereby counteracting the thrust produced.

IV. Propeller Pitch

Along with the diameter, the pitch of the propeller was also changed. Pitch in this case is defined as "The distance which the propeller would screw through the air in one revolution" [9]. The pitch affects how much of the propeller's area is facing the direction of the flow of fluid, therefore changing the drag force experienced as well as the thrust produced (illustrated in Figure 2). This is supported by the claim of David Rogers, stating "Increasing the blade pitch increases the blade drag, while decreasing the blade pitch decreases the blade drag" [11]. Though the blade drag is increased, simultaneously the thrust produced is also increased with a greater surface area accelerating the air, as visible from (5).

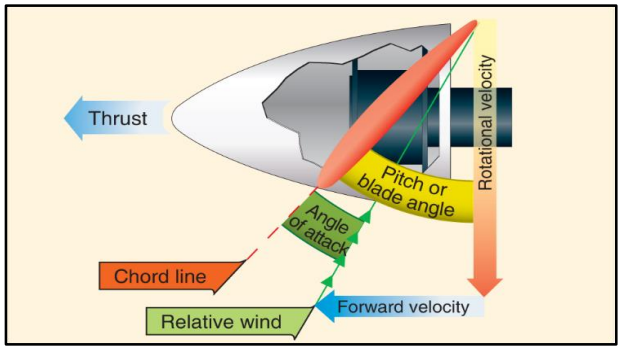


FIGURE 3
DESCRIPTION OF PROPELLER PITCH [10]

LITERATURE VALUES

I. Propeller Pitch vs Efficiency

In order to formulate a hypothesis, literature values showing correlations of pitch and propeller diameter to efficiency was researched. Research conducted by Korukonda Shravan and Swaroop Nagnoori show that increasing the angle of attack at first leads to a linear increase in the lift coefficient, which refers to greater lift produced, followed by a gradual decrease [12].

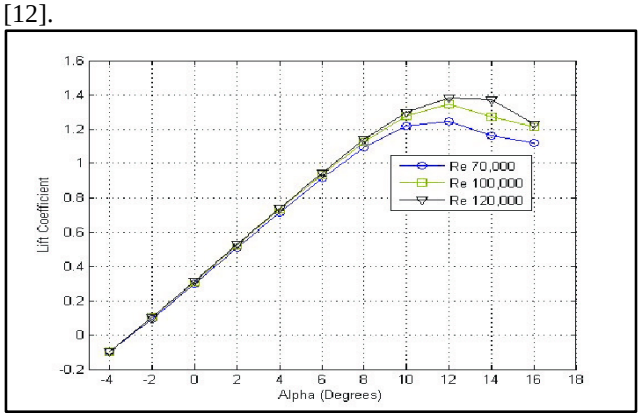


FIGURE 4
LIFT COEFFICIENT VS ANGLE OF ATTACK [12]

The angle of attack is defined as "the angle between the relative airflow and the chord line" [13], and as long as the relative airflow's direction stays the same, an increase in pitch results in an increase in angle of attack, as shown in Figure 3. As such, an increase in pitch results in a slower rate of increase of thrust, hence applying this to (5), the efficiency should follow the same trend as shown in Figure 4. The relationship that describes the curve is a logarithmic function, as the initial slope is steep, but gradually flattens to a near constant value.

II. Propeller Diameter vs Efficiency

The same effect should be seen by increasing the diameter as well, as research conducted by NASA shows near-linear increase in thrust coefficient until it starts to level and

eventually fall with increased blade radius ratio [14]. The blade radius ratio is given as the ratio between the propeller radius (r) and the blade tip radius (R).

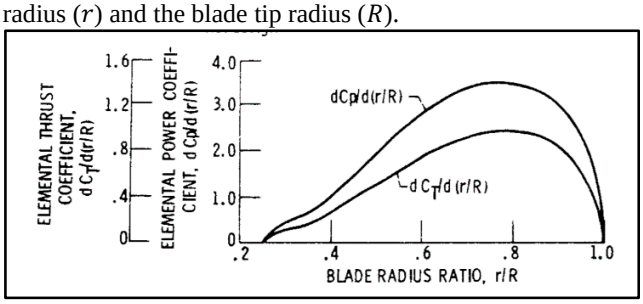


FIGURE 5
BLADE RADIUS RATIO VS THRUST COEFFICIENT [14]

As the blade tip radius will not be altered in this experiment, the value of R can be taken as a constant, and an increase in blade radius will naturally result in an increase in diameter. An increase in the thrust coefficient results in an increase in thrust, by the equation defined below [2]:

$$C_T = T/\rho n^2 D^4 \quad (9)$$

Where n is the rpm of the propeller and D the diameter. Hence again, through (5), the efficiency should increase. Therefore, the relationship shown in Figure 5 should match with a graph plotted against propeller diameter and efficiency, showing linear increase initially, followed by gradual leveling off.

HYPOTHESIS

Considering (5), it is expected that increasing both the pitch and diameter would result in gradual slowing of the rate of increase of overall thrust produced, as greater aircraft velocity increases the drag exponentially. Analyzing the literature values, I predict that efficiency will rise sharply at first, and become constant like a logarithmic curve. The literature values show a sharp decrease after this point, however, due to the small range of propellers being tested, it is unlikely to see this. Though it is hoped that the extent of this relationship can be tested in the experiment, due to only a small range of propellers being available, a complete generalization of the trend may be difficult.

METHODOLOGY

To calculate the efficiency, the thrust and power inputted had to be measured, as well as the RPM for use in analysis. The dimensions of propellers that were tested is shown in Table 1 and 2. Both diameter and pitch were altered by using different sets of propellers manufactured by "Master Airscrew".

First, for calculating thrust, the Vernier force sensor, with uncertainty $\pm 0.001\ N$ was tightly locked into position using a clamp stand. Along with the sensor, a thin block of wood and a 12-volt motor must be attached together. After this, the block of wood is connected to the force sensor using a thin string and placed apart so that there is only slight tension within the string, to ensure the tension measured is all from the thrust instead of the initial tension of the string. The wood must be placed on the same level and directly in front of the sensor to measure the force accurately.

As for determining the power inputted into the motor, the reading of the voltmeter on the lab bench power supply and ammeter was used. The uncertainty of the lab bench power supply is $\pm 0.03\ V$, while the ammeter has uncertainty of $\pm 0.01\ A$. In setting up the ammeter, it was set to detect up to 10 amperes, while the voltage was kept constant at 5 volts.

Finally, the RPM of the propeller was measured using a tachometer with uncertainty of ± 1 RPM for calculation of torque. A reflective tape is placed on the center of the propeller, so that the device could accurately measure the RPM.

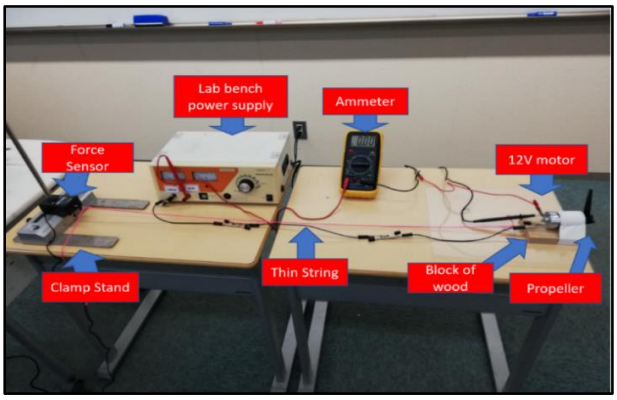


FIGURE 6
ACTUAL PICTURE OF EXPERIMENTAL SETUP USED

To ensure that maximum amount of energy is being converted from electrical power to kinetic energy, the experimental setup was set to minimize energy loss by friction. One material needed to achieve this was the PVC sheet. Research by Fathan Akbar showed that PVC has a low value of coefficient of friction against dry wood, with a value of 0.296 [15], which means that when the block of wood is sliding on the table, less energy is lost as friction.

RAW DATA

Tables 1 and 2 show average values of calculated or measured variables during the experiment from 11 trials. The method used to calculate torque will be explained in the analysis. When testing propeller (E), I realized a weakness in my experimental setup, which was that the large amount of thrust produced by the propeller would push the motor off the ledge. This added extra tension from the gravitational pull that the motor and propeller were experiencing, resulting in greater force measured by the sensor. This is the most likely reason why the thrust generated for propeller (E) was measured to be far greater than that created by propeller (D), with roughly $1.5N$ difference. The same effect could be seen with propeller (D) as well as propeller (I) to a lesser extent. This factor should be considered when evaluating the validity of this experiment.

TABLE I
COLLECTED DATA FOR DIFFERENT PROPELLER DIAMETER

Label	Diameter (m)	Pitch (m)	Avg. Max RPM (±1 RPM)	Avg. Max Thrust (±0.001 N)	Avg. P_{in} (±0.15 W)	Avg. Torque (NM)
A	0.1524	0.1016	8045	0.932	16.55	0.5662
B	0.1778	0.1016	5915	1.291	13.46	0.6874
C	0.2032	0.1016	5492	1.836	16.17	0.7115
D	0.2286	0.1016	5130	2.188	21.35	0.7321
E	0.2540	0.1016	5024	3.716	26.11	0.7381

TABLE II
COLLECTED DATA FOR DIFFERENT PROPELLER PITCH

Label	Diameter (m)	Pitch (m)	Avg. Max RPM (±1 RPM)	Avg. Max Thrust (±0.001 N)	Avg. P_{in} (±0.15 W)	Avg. Torque (NM)
F	0.2032	0.0762	6052	0.618	13.89	0.6796
G	0.2032	0.1270	5118	1.215	25.30	0.6874
H	0.2032	0.1524	5344	2.177	23.15	0.7328
I	0.2032	0.1778	5555	2.118	20.04	0.7199

DATA ANALYSIS

Using the processed data, initially, a graph plotted based on the highest R^2 value was created, the R^2 value, or coefficient of determination, being a number that judges the accuracy of a model [16]. However, both graphs showed a linear trendline, which would suggest that efficiencies greater than 1 could be achieved. This indicates that such trendline is highly unlikely, as it is impossible to obtain values of efficiency greater than 1. Given that, new trendlines were created based on the hypothesis that both diameter and pitch will have a logarithmic relationship. The error bars are based on the standard deviation of the data.

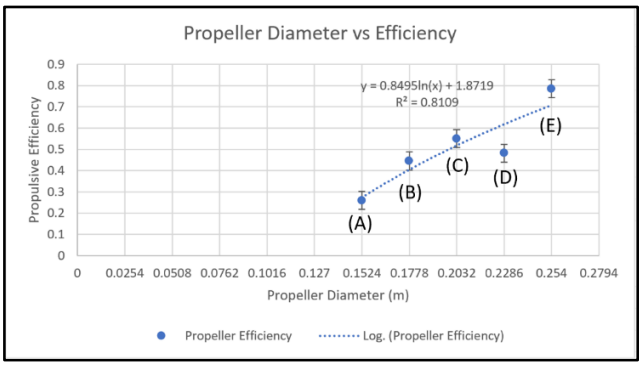


FIGURE 7
TREND OF PROPELLER DIAMETER AND EFFICIENCY BASED ON HYPOTHESIS

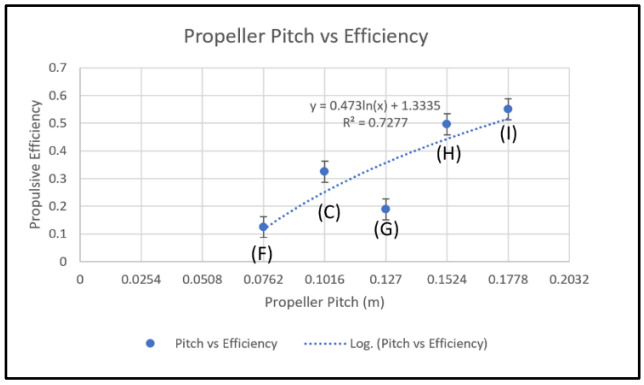


FIGURE 8
TREND OF PROPELLER PITCH AND EFFICIENCY BASED ON HYPOTHESIS

In these two newly created graphs, though the R^2 value decreased, the trend reflects the information found during the background research better, as the rate of increase in efficiency is seen to decrease with greater diameter and pitch. However, these trendlines do not account for the sharp decrease in efficiency for propeller (D) in Figure 7 and propeller (G) in Figure 8. Even considering a slight increase in thrust for propeller (E), it is known to be larger than the rest of the propellers. By this, it does not bring significant enough change in the efficiency to drastically change the overall trend of the graph, hence it cannot explain the low efficiency of propeller (D). The decrease in efficiency for propeller (G) cannot be explained either, for the same reason of the slight increase in thrust of propeller (I) not affecting the overall trend drastically.

Even when considering the fact that literature values showed a sharp decrease in efficiency, they are not followed by a sudden increase immediately after, which makes it likely that these values are anomalous. From the error bars being far from the trendline, the multiple background research suggesting the efficiency curve to be in the shape of a logarithmic graph before dropping, and with nothing strange recorded qualitatively either, it is justified to treat these data points as an outlier. New graphs without the outliers are shown below.

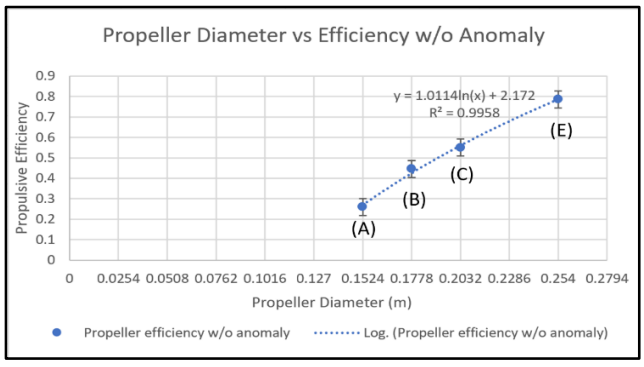


FIGURE 9
TREND OF PROPELLER DIAMETER VS EFFICIENCY EXCLUDING ANOMALY

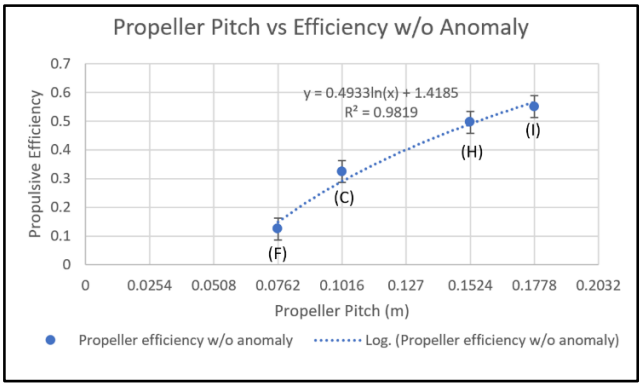


FIGURE 10
TREND OF PROPELLER PITCH VS EFFICIENCY EXCLUDING ANOMALY

With these newly created graphs, my hypothesis seems to be supported by the data, as both best fit trendlines show very high R^2values, showing consistency with the literature. Not only do Figures 9 and 10 indicate a logarithmic relationship, but calculations based on torque and thrust also support the trendline. Frank Hitchens in his book states "The propeller produces the greatest axial thrust for the least amount of engine torque when the maximum thrust/torque ratio is being produced."[13], in other words, a propeller with higher thrust to torque ratio operates at higher efficiency. This ratio can be used as secondary evidence to support the analyzed graphs. Torque was calculated through creating an equation that related the RPM and torque based on information given by the producer of the motor.

Tables III and IV lists the thrust-torque ratios for each propeller. Due to the ratio being reliant on thrust, it is unreliable for propellers (D), (E), and (I) from the fact that the measured thrust was larger, as stated in the raw data section. However, the general trend could be analyzed.

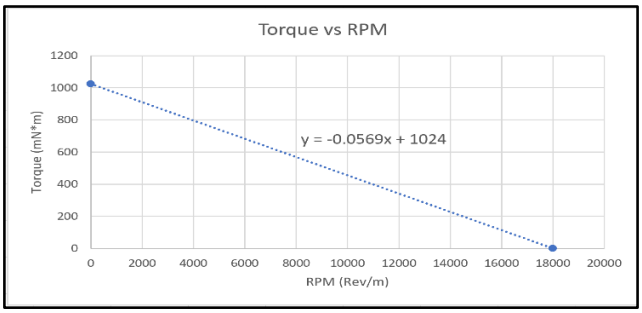


FIGURE11
GRAPH WHERE THE SLOPE OF THE GRAPH CALCULATES TORQUE

TABLE III
THRUST TORQUE RATIO FOR DIFFERING PROPELLER DIAMETERS

Propeller Diameter (m)	Max Thrust-Torque Ratio (±0.01)
0.1524	1.65
0.1778	1.88
0.2032	2.58
0.2286	2.99
0.2540	5.03

TABLE IV
THRUST TORQUE RATIO FOR DIFFERING PROPELLER PITCH

Propeller Pitch (m)	Max Thrust-Torque Ratio (±0.01)
0.0762	0.91
0.1016	1.88
0.1270	1.66
0.1524	3.02
0.1778	2.99

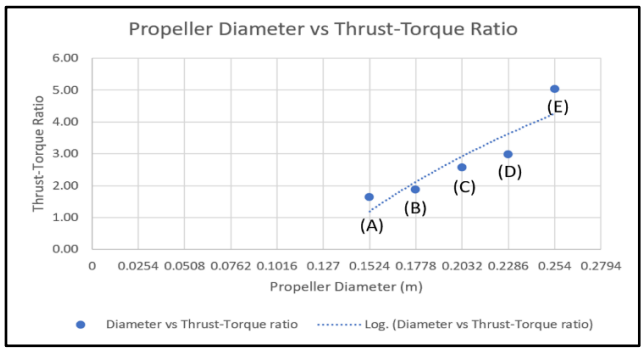


FIGURE 12
GRAPH OF PROPELLER DIAMETER VS THRUST TORQUE

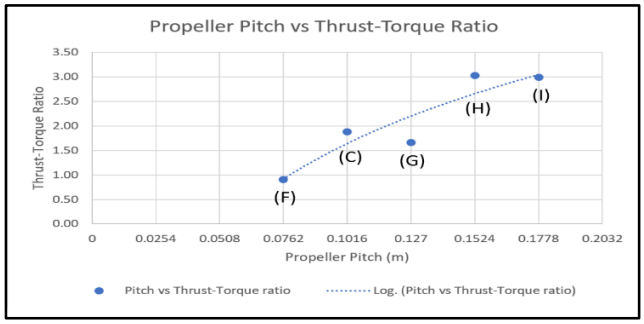


FIGURE 13
GRAPH OF PROPELLER PITCH VS THRUST TORQUE

Considering that Figure 12 and 13 show the value of anomalous and unreliable data, the results are still very close to the result shown in Figure 9 and 10. The trend of thrust to torque ratio, which measures efficiency, matching Figure 9 and 10 provide further support that the trend analyzed is correct. By this, the relationship being logarithmic is further supported from the data above.

RELIABILITY OF RESULTS

Although the hypothesis was supported without the anomaly, there are limitations to the reliability of Figures 9 and 10, as an entire data point had to be removed. The removal can be justified from the fact that the data point was far from the overall trend, and no literature or qualitative observations exist to support such phenomenon, but my collection of data does not go far enough to know whether there are other potential drops in efficiency outside of the small scope tested. To strengthen the analysis of the data, there needs to be further testing of wider ranges of propellers to determine whether the data point removed could be considered an anomaly. One possible reason for such anomaly could be due to the propellers having a fixed pitch, which means they

operate efficiently only during specific conditions of RPM and airspeed [17]. Both anomalies had very similar average RPM values of 5118 and 5130, however it is uncertain whether this factor was the cause of the inefficiency.

Furthermore, there are also flaws within the experiment itself that weakens the validity of the analysis. One source of this flaw is the setup of calculating thrust, combined with the vibration of the motor. As mentioned in the raw data section, for propellers (D), (E), and (I), the motor fell off the table, causing an increase in the thrust measured. Trying to solve this problem resulted in the vibration of the motor to shift the entire data collection system, making the data collected unreliable. To solve this issue, a frame designed specifically for the motor should have been used to minimize the effect of the vibration.

However, despite these flaws that lessen the confidence about the relationship between the propeller pitch and diameter against propulsive efficiency, the overall logarithmic trends of the analyzed data can be considered as valid as they agree with the background research. In the case of propeller diameter, NASA showed larger diameter propellers being more efficient [14], and with my experiment, it showed the same trend. As for the pitch, the conclusion reached by Korukonda Shravan and Swaroop Nagnoori in their research [12] show very similar trendline to the one achieved in this experiment.

CONCLUSION

From the data analysis, I conclude that there is an increase in the propulsive efficiency as both propeller diameter and pitch is increased, and this increase is likely logarithmic. This conclusion is supported by both accepted literatures as well as the overall conclusion drawn from the analysis of the processed data. Though there were limitations within the data collected that makes generalization and reliability of this relationship weak, within the scope of my experiment, it can be concluded that pitch and diameter have a positive logarithmic relationship with propulsive efficiency. By this, the most efficient propeller design for drones flying under hover condition are those with a large diameter and pitch.

ACKNOWLEDGMENTS

I wish to thank Mr. Mackenzie, Mr. Zara and Mr. Sutton for their help and guidance in conducting the experiments and proofreading this paper.

REFERENCES

[1] "Why Are Drone Flight Times so Short - How to Improve Your Flight Time." *Zone Science*, 7 Dec. 2018, zonescience.com/why-are-drone-flight-times-so-short/.

[2] Spakovszky, Zoltan S. "16. Unified: Thermodynamics and Propulsion ." Edited by D Quattrochi, *11.7 Performance of Propellers*, MIT, 1 Feb. 2002, web.mit.edu/16.unified/www/SPRING/propulsion/notes/node86.html.

[3] Theys, Bart, et al. *ResearchGate*. 2016, *ResearchGate*, www.researchgate.net/publication/304817071_Influence_of_propeller_configuration_on_propulsion_system_efficiency_of_multi-rotor_Unmanned_Aerial_Vehicles.

[4] Brezina, Aron J, and Scott K Thomas. Measurement of Static and Dynamic Performance Characteristics of Small Electric Propulsion Systems. Measurement of Static and Dynamic Performance Characteristics of Small Electric Propulsion Systems, web1.cs.wright.edu/~sthomas/aronsconferencepaper.pdf.

[5] Hall, Nancy. "Equation of State." *NASA*, NASA, 5 May 2015, www.grc.nasa.gov/www/k-12/airplane/eqstat.html.

[6] Hall, Nancy. "Propeller Thrust." *NASA*, NASA, 5 May 2015, www.grc.nasa.gov/www/k-12/airplane/propth.html.

[7] OpenStaxCollege. "Drag Forces." *College Physics*, 23 Jan. 2012, opentextbc.ca/physicstestbook2/chapter/drag-forces/#import-auto-id1165298787035.

[8] Vorkoetter, Stefan. "Propeller Basics." *Http://Www.stefanv.com*, 1 Mar. 2002, www.stefanv.com/rcstuff/qf200203.html.

[9] FAATest. *Basic Propeller Principles*, Dauntless Software, faatest.com/books/FLT/Chapter17/BasicPropellerPrinciples.htm.

[10] "Propeller Blade Angle." *Aerodynamics: Propeller Basics* , 5 July 2016, learntoflyblog.com/2016/07/05/aerodynamics-propeller-basics/.

[11] Rogers, David F. American Bonanza Society, 2010, www.nar-associates.com/technical-flying/propeller/cruise_propeller_efficiency_screen.pdf.

[12] Korukonda, Shravan, and Swaroop Nagnoori. Numerical Investigation of an Airfoil at Low Reynolds Numbers for MAV Application. ResearchGate, 2015, Numerical Investigation of an Airfoil at Low Reynolds Numbers for MAV Application, www.researchgate.net/publication/280931288_Numerical_Investigation_of_an_Airfoil_at_Low_Reynolds_Numbers_for_MAV_Application.

[13] Hitchens, Frank E. Propeller Aerodynamics The History, Aerodynamics & Operation of Aircraft Propellers. Andrews UK Limited, 2015.

[14] Mikkelson, Daniel C, et al. Design and Performance of Energy Efficient Propellers for Mach 0.8 Cruise. NASA, 1977, Design and Performance of Energy Efficient Propellers for Mach 0.8 Cruise, ntrs.nasa.gov/search.jsp?R=19770013165.

[15] Akbar, Fathan, et al. Minimal Tools for Accurately Measuring the Coefficient of Kinetic Friction. 2019, Minimal Tools for Accurately Measuring the Coefficient of Kinetic Friction, www.researchgate.net/publication/331034525_Minimal_Tools_for_Accurately_Measuring_the_Coefficient_of_Kinetic_Friction.

[16] Grant, Mitchell, and Will Kenton. "How the Coefficient of Determination Works." *Investopedia*, Investopedia, 24 Apr. 2019, www.investopedia.com/terms/c/coefficient-of-determination.asp.

[17] Federal Aviation Administration. *Pilot's Handbook of Aeronautical Knowledge*. Skyhorse Pub., 2014.

Types of Healthcare Systems and the United States Healthcare

Anna Toneva
Prospect High School, Mount Prospect, United States
annatoneva@outlook.com

Abstract - **The United States is the only developed country that leaves a large part of its population without healthcare coverage. There are four main models of healthcare and the developed countries tend to follow one of three. However, the United States doesn't follow any of them. The United States currently doesn't have a set healthcare system, but rather has private insurers with their own sets of rules. There are many complaints about the current American healthcare system and people are pushing for change. Most U.S citizens would receive a healthcare plan from their employers if they work full time. If a citizen is unemployed or they make below a certain amount of money they might qualify for Medicaid.**

Key Words - funding, healthcare, medicaid, taxes

Introduction

Would one healthcare system be able to accommodate the United States? The American Medical Association defines health care as a fundamental human good because it affects peoples opportunity to pursue life goals, reduces pain and suffering, helps prevent premature loss of life, and provides information needed to plan for people's lives healthcare is to maintain population health, treat those who are sick, and protect families from financial ruin [14]. Around the world, there are four main models of healthcare implemented. Those models are The Beveridge Model, The Bismarck Model, The National Health Insurance Model (NHI), and The Out-Of-Pocket Model [2][10][12]. These models each have advantages and disadvantages, some more than others. Certain healthcare models are only able to be implemented in countries that are more developed because they would be unsustainable in less developed countries [15].

System Comparison

The Beveridge Model was named after a social reformer, William Beveridge. Sir William Beveridge was a well-known and respected educator and civil servant. After world war 2, in the summer of 1941, he was asked to devise a plan of social reconstruction. Sir William Beveridge, in the past, had worked on the poor side of the East End of London. While he was working he saw what kind of medical service, child support, and housing was present there and how it was run with different rules by different operators [17]. It consists of the government being the only payer for health care. Having the government as the sole payer keeps costs low, and it standardizes benefits across the country. The funding for the Beveridge model comes from income taxes. This ensures that every person living in a country that has the Beveridge Model implemented has access to healthcare [10]. Italy for example, has the Beveridge Model implemented and all citizens, legal or non-legal, have access to free or low cost access to the healthcare system [13]. Other countries that use the Beveridge Model include Denmark, Italy, Finland, Britain, and New Zealand [1].

The Bismarck Model was created by Otto von Bismarck in the late 19th century. The Bismarck Model is used by countries such as Belgium, Switzerland, Japan, Germany, and the Netherlands [1]. This model primarily funds healthcare through payroll deductions, from employers and employees, via holding [5][10][12]. The Bismarck Model ensures that everyone has access to healthcare and mitigates the insurance industry's opportunity to make a profit. The model has survived the militarism of Germany during World War 1, Nazism, World War 2, and the aftermath of World War 2 [5].The model

surviving these events proves that it is an effective and durable form of healthcare. France has the Bismarck Model implemented, and it's mainly funded by payroll taxes (50%), followed by national earnmark taxes (35%), voluntary health insurance companies (13%), and state subsidies (2%) [4].

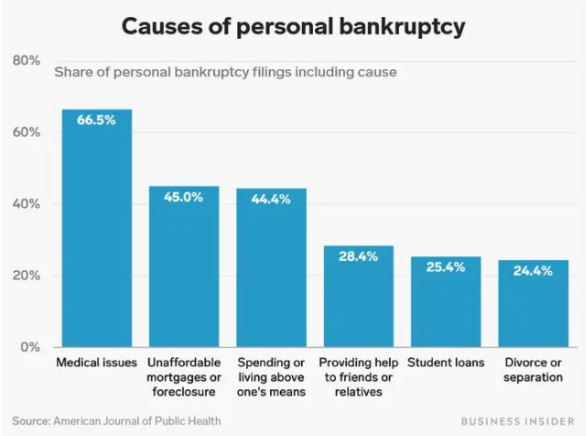
The National Health Insurance Model (NHI) has commonalities to the Beveridge and Bismarck Model. Similar to the Beveridge Model the government pays for the healthcare with taxes, and similar to the Bismarck model, the providers are private [6][8]. Canada, for example, funds their healthcare through taxation and public funds [3]. The mix of private practice and public insurance allows hospitals to have independence while minimizing insurance company complications [6][8]. For the most part, patients are able to choose the provider for their healthcare; however, there are some requirements Canadians have to follow to be eligible for their healthcare. For instance Ontario residents are required to be in the province at least 153 of the last 183 days. Being out of the country for more than 30 days, prevents automatic eligibility for the benefits of healthcare until the requirement is met [3]. The National Insurance Model is also used by countries such as South Korea and Taiwan [1].

The Out-of-Pocket Model is mainly used in less developed countries. Patients are expected to pay for their procedure on their own without any aid from the government. This system severely limits who can receive medical attention. Those who can not afford their procedure and medication are unable to receive it. Such a system benefits the wealthy and leaves the impoverished with little to no access to care due to the high costs of treatment. Low-income individuals may rely on the relief efforts of non-government organizations [16].The Out-of-Pocket Model is used by countries such as rural parts of China, India, parts of Africa, and parts of South America [1].

The United States' Current Position

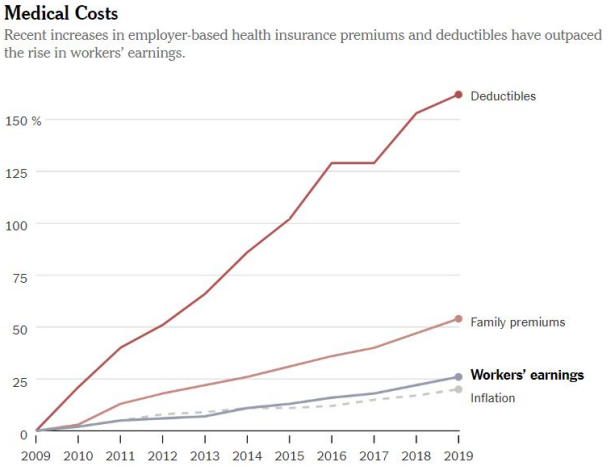
Is it realistic to have just one healthcare model implemented in the United States? The main issue the United States is likely to face if they were to implement a new healthcare system is choosing one that is beneficial to the general population. The healthcare that is currently implemented has many issues that people voice clearly: cost, access, and quality. A big issue is that people are expected to pay for their own expensive procedures and medications with the help of a private healthcare insurance [9]. Most people that have a private healthcare insurance receive it through their employer. If a United States citizen is uninsured and make $29,435 or less a year for a family of three or $17,263 or less as an individual a year, they may qualify for Medicaid in 32 states. Coverage varies depending on which state a person lives [7][11]. Medicaid is a healthcare system that helps low income families with hospital stays, individual doctors visits, long-term medical conditions, and more. Medicaid is funded by the federal government and run at the state level. [7][11].

According to the American Journal of Public Health, lately, cost increases have come in the form of high deductibles. This shifts the system closer to the Out of Pocket Modeland is blamed for increase in personal bankruptcies[19].



Employer based coverage have declined as costs have increased. According to the Kaiser Family Foundation this decline was from 60% of

the nonelderly population in 2008 to 55% in 2013 [18].



How The Models Would Affect the United States

If the United States were to implement the Beveridge Model, that would likely mean taxes would be raised. The Bismarck Model, if implemented in the United States, would be funded by payroll deductions. The rest of the funding would be provided through national earnmark taxes, state subsidies, and voluntary health insurance. If the National Health Insurance Model were to be implemented in the United States there would be more room for the government to control the health system while every person who abides by the rules of the system will have access to healthcare. If the United States were to implement the Out-of-Pocket Model, everyone would pay for themselves and there would be no insurance providers. Having an Out-of-Pocket Model in the United States would leave patients very susceptible to bankruptcy.

Conclusion

If recent trends continue the cost of employer based insurance could become unaffordable and the number of people covered in this way would continue to decline. This will shift more of the cost to public payers and may be the catalyst needed to push the United States in the direction of system that covers all. If the United States if forced in the direction of publicly funded healthcare system, it is likely to follow the National Health Insurance Model. This system works in Canada and given the cultural and economic similarities between the two countries it is likely to work in the USA. A big problem the United States' healthcare currently has is funding for medical procedures. The National Health Insurance Model has funding from the government and the providers are private. The National Health Insurance model seems the most reasonable healthcare model to implement in the United States. The United States would still be able to have private providers but the government would fund the healthcare through taxes.

Acknowledgments

The author wishes to thank Ms. Lori Amedeo for her guidance and support throughout the writing of the paper, Mr. Paul Lewis for introducing the conference and his support, and last, but not least, the authors father for his constant support and encouragement.

References

[1] Brown, Dave. *Health Care Systems: Differences Around the Globe*. 02 Jan.. 2014, www.globalpremeds.com/blog/2014/02/22/health-care-systems-differences-around-the-globe/.

[2] Chung, Mimi. "Health Care Reform: Learning From Other Major Health Care Systems | Princeton Public Health Review." *Princeton University*, The Trustees of Princeton University, 2017, pphr.princeton.edu/2017/12/02/unhealthy-health-care-a-cursory-overview-of-major-health-care-systems/.

[3] Doughlas, Tommy. "Understanding The Canadian Health Care System - Health Care in Canada." *International Travel Insurance Group*, 2004, www.internationalinsurance.com/health/systems/canadian-health-care.php.

[4] Durand, Isabelle. "International Health Care System Profiles." *France : International Health Care System Profiles*, 2017, international.commonwealthfund.org/countries/france/.

[5] K., K. "The Bismarck Model." *The Bismarck Model*, 2 Jan. 2011, healthmatters4.blogspot.com/2011/01/bismarck-model.html.

[6] K., K. "The National Health Insurance Model." *The National Health Insurance Model*, 11 Jan. 2011, healthmatters4.blogspot.com/2011/01/national-health-insurance-model.html.

[7] Kagan, Julia. "Medicaid." *Investopedia*, Investopedia, 12 Mar. 2019, www.investopedia.com/terms/m/medicaid.asp.

[8] Nguyen, Angeline. "International Health Care Systems Part 3: The Bismarck Model." *Morning Sign Out*, 5 Aug. 2017, morningsignout.com/international-health-care-systems-part-3-the-bismarck-model/.

[9] "Problems With Health Care in the United States." *Problems of Health Care in the United States*, 2011, saylordotorg.github.io/text_social-problems-continuity-and-change/s16-04-problems-of-health-care-in-the.html.

[10] Reid, T.R. *Health Care Systems - Four Basic Models*. 2009, www.pnhp.org/single_payer_resources/health_care_systems_four_basic_models.php..

[11] Tricia, Brooks. "Where Are States Today? Medicaid and CHIP Eligibility Levels for Children, Pregnant Women, and Adults." *The Henry J. Kaiser Family Foundation*, 1 Apr. 2019, www.kff.org/medicaid/fact-sheet/where-are-states-today-medicaid-and-chip/.

[12] Wallace, Lorraine S. "A View Of Health Care Around The World." *The Annals of Family Medicine*, 2012, www.annfammed.org/content/11/1/84.1.full.

[13] Winke, Rebecca. "Healthcare in Italy: What You Need to Know." *ITALY Magazine*, 17 Jan. 2019, www.italymagazine.com/featured-story/healthcare-italy-what-you-need-know.

[14] "Defining Basic Health Care." *American Medical Association*, www.ama-assn.org/delivering-care/ethics/defining-basic-health-care.

[15] Karamat, Jawad, et al. "Promoting Healthcare Sustainability in Developing Countries: Analysis of Knowledge Management Drivers in Public and Private Hospitals of Pakistan." *International Journal of Environmental Research and Public Health*, MDPI, 12 Feb. 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6388157/.

[16] Partners, Zanmi. "Haiti- Partners In Health." *Partners In Health*, www.pih.org/country/haiti.

[17] Klein R. *The New Politics of the NHS.* 3rd ed. New York, NY: Longman; 1995.

[18] "Employer-Sponsored Coverage Rates for the Nonelderly by Age." *The Henry J. Kaiser Family Foundation*, 29 Nov. 2018, www.kff.org/other/state-indicator/rate-by-age-2/?dataView=0&activeTab=graph¤tTimeframe=0&startTimeframe=9&selectedDistributions=total&selectedRows=%7B%22wrapups%22%3A%7B%22united-states%22%3A%7B%7D%7D%7D&sortModel=%7B%22colId%22%3A%22Location%22%2C%22sort%22%3A%22asc%22%7D.

[19] Hoffower, Hillary. "Staggering Medical Bills Are the Biggest Driver of Personal Bankruptcies in the US. Here's What You Need to Know If You're Thinking about Filing for Bankruptcy." *Business Insider*, Business Insider, 25 June 2019, www.businessinsider.com/causes-personal-bankruptcy-medical-bills-mortgages-student-loan-debt-2019-6.

The Short-Term Effects of Sugar on Self-Worth

Emmalyn Kartchner
Yokota High School, Fussa, Japan
Email: ekar8954@student.dodea.edu

***Abstract* – As all forms of sugar become increasingly prevalent globally, it becomes a necessity to understand the consequences of a high sugar intake, not only on the human body but also on the brain. The current study observes the short-term correlation of refined sugar on feelings of self-worth in teenagers. The Rosenberg Self-Esteem Scale was used to assess the high-school-aged participants' levels of self-worth before and after administration of either a sugar, sugarless, or no cookie, based on randomly generated grouping software. RSES results showed that the average score of the students who ate the sugar or sugarless cookie decreased by 0.1154 and 0.7083 respectively, while control group's average before and after scores were identical. Consistent with the hypothesis, the comparison of the averages of each group's before and after scores determined that sugar does seem to impact levels of self-worth an hour after consumption in healthy teenagers. However, the data indicated that other refined carbohydrates also have similar effects, suggesting that foods with a high Glycemic Index affect levels of self-worth.**

Key Words – (Refined Sugar, Rosenberg Self-Esteem Scale, Self-Worth, Short-Term Effects, Sugar Cookie)

INTRODUCTION

The purpose of this experiment was to test the effect, if any, of refined sugar on short term self-worth in teenagers. This was done by conducting an experiment using groups of high schoolers: those who were given a homemade cookie containing refined sugar, those given a homemade cookie without sugar, and those who were given nothing. Short term is characterized as the time whereupon sugar has been absorbed into the bloodstream and circulates in the body, while self-worth is an individual's subjective evaluation of their own worth, which heavily influences their choices and decisions.

Studies indicate that the levels of depression seem to have increased dramatically over recent years, raising the question of its correlation with the growing levels of sugar in food substances. Noticing from research and personal experience that eating large amounts of refined carbohydrates including sugar results in "sugar crashes" which encompass mood swings and depression (feelings of severe despondency and dejection), I wondered whether or not sugar can also affect mental health and specifically feelings of self-worth, in the short term, when smaller but still excess amounts of sugar are consumed. Many can also relate to the common experience of having self-esteem fluctuate throughout the day rather than stay fixed. This observation introduces a margin of experimentation which could take place, on whether or not sugar consumption influences these self-esteem fluctuations in the short term, as well as the large scale "sugar crashes" that can cause mood swings and depression.

As refined sugar and other artificial sweeteners become increasingly more prevalent in not only sweets, but snacks, drinks, and even ingredients in every-day foods, it is imperative to understand its consequences, not only on physical health, but also mental health. While many people have experienced and noticed the effects of sugar on their body and mood, many still consume it regularly because the consequences are not completely apparent or significant enough to limit intake. If hard evidence can be found to suggest the effect of refined sugar on self-worth, the public may raise their awareness or desire to limit their own intake, and this could eventually lead to governments limiting the amount of refined sugar permitted in food substances which would greatly improve the health of its citizens. Although the effects of refined sugar on self-worth may not prove to be vastly different, even pinpointing small changes may lead to explanations of the larger and long-term consequences such as depression.

LITERATURE REVIEW

"Sugar" is a word that often sparks automatic excitement in any human to whom it is spoken to. Upon consumption, sugar also creates a high in the body, as the brain releases dopamine and serotonin which boost mood and stimulate the nucleus accumbens (the area of the brain associated with reward [1]. While these neural systems evolved to reinforce and motivate food intake and foraging for the welfare of the body, they now become out of hand in a world where there is a multitude of readily available food containing extreme amounts of sugar that humans' bodies are not designed to consume [2].

To give context, the highly recommended intake of sugar, according to the American Heart Association which aligns with worldwide expert panel recommendations including the World Health Organization, is 25 grams for women and 37.5 grams for men or 100 to 150 calories

respectively [3]. However, one 12-ounce can of soda has about 39 grams of sugar which is not only more than the recommended intake but also is usually consumed at once, meaning that the blood sugar rapidly rises to an extreme amount. To give another example, one serving of Chips Ahoy! Cookies (3 cookies) contains 11 grams of sugar. Although this is not over the recommended daily intake, consuming ⅓ of recommended daily intake of sugar in a condensed food rather than dispersed throughout the natural diet, also spikes blood sugar which then drops lower soon after.

As described in article [4], "Self-esteem begins with our bodies. Since mind and body are one entity, the smooth, interrelated functioning of our body parts and our brain chemistry provide the foundation for an inherent sense of wellness." While it is heavily researched that sugar causes chronic diseases including hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease [5]-[7], the link between self-esteem and sugar is proposed but not heavily researched. What is currently known is that sugar has some type of correlation with long term cognitive decline such as illustrated in the research paper titled, "The emerging role of dietary fructose in obesity and cognitive decline" [7] which explains that high-fructose corn syrup accounts for as much as 40% of caloric sweeteners used in the United States and may be related to the increase in Alzheimer's disease.

Studies such as "Immediate effects of chocolate on experimentally induced mood states" [8], in 2007, have demonstrated sugar's immediate benefit on lab induced negative mood. The authors ascribed the 3-minute improvement of mood after eating a piece of chocolate, to the palatability of this particular sweet food. While this study examines immediate effects, it does not examine short term effects on mood such as due to levels of blood-sugar (which are affected by refined sugar since carbohydrates that do not contain fiber, easily cause high blood sugar). Other studies have been carried out such as "The Effects of Self-Esteem and Stress on Eating Behaviours in Females" by Ashley Albrecht in 2014 [9] that address short term correlation between sugary foods and self-esteem and stress, but in a different light.

There seems to be a strong relation between low self-esteem and depression, but as stated in the paper, "Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies" by Sowislo, J.F., and Orth, "there is not yet consistent evidence on the nature of the relation." [10] They further describe that "the vulnerability model states that low self-esteem contributes to depression while the scar model states that depression erodes self-esteem." Filling in prior research on the subject, by experimenting the short-term effect of refined sugar on self-esteem, may help to conclude whether low self-esteem contributes to depression, or vice versa. These findings would contribute to doctors' treatment of patients with depression, and prompt interventions, using knowledge of refined sugar's effect, aimed at increasing self-esteem to reduce the risk of depression.

HYPOTHESIS

The purpose of this study is to examine the effects of refined sugar on the short-term levels of self-worth in teenagers. Using knowledge based on previous studies, it is anticipated that teenagers will score lower on the Rosenberg Self-Esteem Scale [11] one hour after consuming the cookie containing refined sugar, while the average score on the RSES of the control group and consumers of the sugarless cookie will show no significant change in self-esteem.

VARIABLES

- **Independent Variable**: The independent variable in this investigation was the refined cane sugar in the consumed cookie. For a third of the teenagers who participated, the cookie contained sugar, and another third of the teenagers had the sugarless cookie.
- **Dependent Variable:** The dependent variable in this experiment was the score on the test using the Rosenberg Self Esteem Scale [11].
- **Control Variable:** The participants randomly assigned to the control group did not consume a cookie.

METHODOLOGY

Trial Participants

In total, the sample will consist of 40 participants (17 female, 23 male) between ages of 14 to 18 (mean: 15.925, standard deviation: 0.917) taken from a local high school. The participants were students in the classes whose teachers volunteered a section of their class time for this experiment on the two designated days. Before the experiment, all participants were asked if they agreed to take a short quiz and if they had any health concerns that would not allow them to eat a cookie (containing butter, eggs, vanilla extract, all-purpose flour, baking powder, and salt). All the students who were asked to participate, agreed, on account that they did not have any medical conditions and felt comfortable in the situation.

Characteristics	(N=40)
Variable	**Frequency (%) or mean (SD)**
Sugarless Cookie Participants	**(N=12)**
Age (mean years)	15.83 (±0.937)
Gender	
Male	7 (58.3)
Female	5 (41.7)
Academic Level	
9	1 (8.3)
10	3 (25.0)
11	7 (58.3)
12	1 (8.33)
Sugar Cookie Participants	**(N=13)**
Age (mean years)	15.77 (±0.927)
Gender	
Male	8 (61.5)
Female	5 (38.5)
Academic Level	
9	3 (23.1)
10	2 (15.4)
11	8 (61.5)
12	0 (0)
Participants with no cookie	**(N=15)**
Age (mean years)	16.13 (±0.915)
Gender	
Male	8 (53.3)
Female	7 (46.7)
Academic Level	
9	1 (6.67)
10	3 (20)
11	2 (13.3)
12	9 (60)
SD: Standard Deviation	

FIGURE 1
BASIC CHARACTERISTICS TABLE OF PARTICIPANTS, INCLUDING AGE, ACADEMIC LEVEL (GRADE), AND GENDER

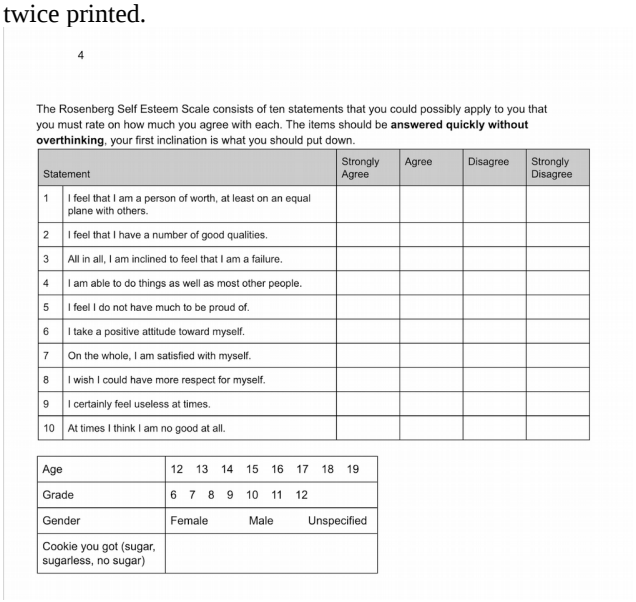
Trial Design

The Rosenberg Self-Esteem Scale [11] was administered as a paper copy, to participants. This is a 10-item scale that measures global self-worth by evaluating both positive and negative feelings about the self. Participants responded using a 4-point Likert scale that ranges from strongly agree to strongly disagree. The test consists of questions such as, "I feel that I have a number of good qualities" and "all in all, I am inclined to feel that I am a failure". This survey instrument has demonstrated a high degree of reliability and validity. Test-retest correlations of the Rosenberg scale are said to range from .82 to .88 with a coefficient alpha ranging from .77 to .88, indicating a high degree of internal consistency. The scale is also found to have adequate convergent validity as it was closely associated with similar measures (e.g. the Coopersmith Self-Esteem inventory). [9]

A paper copy of this test will be administered to all participants (students who consented to be tested) directly before, and an hour after the food is given. The before assessments were numbered one to forty, as were the after assessments, and each student was given the same number both times. This way, groupings were kept track of without disclosing personal information. To randomly divide students into groups (sugar cookie, sugarless cookie, no cookie), the participants names were inputted into random grouping software, which matched each name to a group. Based on these matchings, a student was given the corresponding cookie. Each classroom included students in all three groups, to reduce bias based on the environment of individual classes. It was unnecessary to conceal the conditions of other groups, since it was mentioned to all participants that the groupings were random and had nothing to do with individuals, therefore eliminating the interference of the participants assuming favor, on self-esteem.

Materials

Two types of cookies were home baked for this experiment. All cookies contained equivalent amounts of butter, eggs, vanilla extract, all-purpose flour, baking powder, and salt. White sugar was then included in only half of the cookies, introducing the independent variable. The sugar was omitted from the recipe so the other half of cookies did not contain the sugar. Two paper printouts of the Rosenberg Self Esteem [11] were given to each participant, both with the same number on them. Thus, 80 copies of the RSES assessment were made, with numberings from one to forty twice printed.



4

The Rosenberg Self Esteem Scale consists of ten statements that you could possibly apply to you that you must rate on how much you agree with each. The items should be **answered quickly without overthinking**, your first inclination is what you should put down.

	Statement	Strongly Agree	Agree	Disagree	Strongly Disagree
1	I feel that I am a person of worth, at least on an equal plane with others.				
2	I feel that I have a number of good qualities.				
3	All in all, I am inclined to feel that I am a failure.				
4	I am able to do things as well as most other people.				
5	I feel I do not have much to be proud of.				
6	I take a positive attitude toward myself.				
7	On the whole, I am satisfied with myself.				
8	I wish I could have more respect for myself.				
9	I certainly feel useless at times.				
10	At times I think I am no good at all.				

Age	12 13 14 15 16 17 18 19
Grade	6 7 8 9 10 11 12
Gender	Female Male Unspecified
Cookie you got (sugar, sugarless, no sugar)	

FIGURE 2
EXAMPLE OF THE RSES ASSESSMENT GIVEN TO EACH PARTICIPANT (TWICE)

Procedure

Sugar cookies were made the night before the experiment, following the recipe titled The Best Rolled Sugar Cookies, by Jill Saunders, on Allrecipies.com [12]. Randomly generated grouping of names was also done ahead of time, and recorded to use as reference when passing out cookies during the experiment. At the designated time, high school students were seated in their classrooms, and were then asked if they agreed to participate after explaining that the experiment included taking a short assessment on their levels of self-esteem as well as possibly eating a homemade cookie. It was ensured that each participant knew that grouping was random and that specific information about the experiment was withheld, such as the retesting at the end, to prevent bias. One assessment using the RSES was passed out to each student, passed from right to left and row by row of desks starting with the front, to ensure that each participant would get the same numbered sheet so that

results for each student could later be paired. The students were then instructed to begin the assessment and were given two minutes to complete it. After the students turned their assessments in to the administrator, the cookies were passed out based on the recorded randomly generated grouping, and students were instructed to eat the entire cookie or in the case of the control group, wait. Once these proceedings were completed, the participants were let alone to engage in whatever was happening in the class, and the administrator kept track of the time since all participants with a cookie had finished eating. After an hour, the RSES assessment was administered to the students again in the same manner, making certain that each student got the same numbered sheet. This time, the assessment also asked for information on grade, age, gender, and what group (sugar, sugarless, nothing) they were in during the experiment. Later, the RSES tests were assessed based on their answers using the 4-point Likert scale, as described in the next section of this paper.

Statistical Analysis

There were 41 students tested. 13 consumed the sugarless cookie, 13 consumed the sugar cookie, and 15 students had nothing. The data was analyzed using averages so that the 2-person discrepancy in sample size did not create a significant influence. The data was analyzed using a point system where, questions 1,2,4,6, and 7 are given three points for strongly agree, two for agree, one for disagree, and 0 for strongly disagree, and questions 3,5,8,9,10 (which are reversed in valence) are given 0 points for strongly agree and up from there. The total score of each test was calculated using the point system, and the averages of all student's before results were calculated, as well as for the after results.

ANALYSIS

	before score	after score	Age	Grade	Gender
No Sugar	16	15	15	10	Female
No Sugar	17	14	14	9	Male
No Sugar	15	15	15	10	Female
No Sugar	19	16	15	10	Female
No Sugar	23	23	16	11	Male
No Sugar	23	24	16	11	Female
No Sugar	15	15	17	11	Male
No Sugar	23	23	16	11	Male
No Sugar	6.5	3	16	11	Female
No Sugar	19	19	17	12	Male
No Sugar	30	30	17	11	Male
No Sugar	23	24	16	11	Male
Sugar	25	24	17	11	Female
Sugar	25	19	17	11	Female
Sugar	13	14	15	10	Male
Sugar	16	16	15	9	Male
Sugar	15	14	14	9	Male
Sugar	23	27	15	9	Male
Sugar	21	20	15	10	Male
Sugar	23	20	17	11	Female
Sugar	12	13	16	11	Female
Sugar	30	30	16	11	Male
Sugar	25	25.5	16	11	Male
Sugar	16	16	16	11	Male
Sugar	19	23	16	11	Female
Nothing	11	12	16	11	Male
Nothing	13	11	17	11	Female
Nothing	15	14	16	11	Male
Nothing	15	12	15	10	Female
Nothing	27	25	16	11	Female
Nothing	30	30	17	12	Male
Nothing	14	12	16	10	Male
Nothing	7.5	8.5	14	9	Female
Nothing	25	25	17	12	Male
Nothing	25	26	15	10	Male
Nothing	14	15	16	11	Female
Nothing	25	28	16	11	Male
Nothing	20	23	17	11	Female
Nothing	18	19	17	11	Female
Nothing	14	13	17	11	Male
Total Average	766/40 = 19.15	756/40 = 18.9	637 / 40 = 15.92	425 / 40 = 10.625	

FIGURE 3
COMPILED DATA TABLE INCLUDING COOKIE EATEN, BEFORE AND AFTER SCORE OF PARTICIPANTS, AND BASELINE CHARACTERISTIC

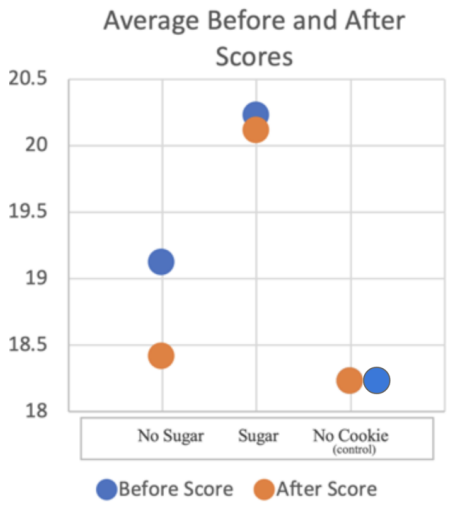


FIGURE 3

Figure 3 compares the before and after scores of the groups who ate the no-sugar cookie, the sugar-cookie, and the control. The averages of each groups before scores, as shown in blue, were as follows: sugarless cookie participants scored on average 19.125, for the sugar cookie group it was 20.23076923, and the control group's average

before score was 18.2333. The averages of each groups after scores, as shown in red were as follows: sugarless cookie participants scored on average 18.416, for the sugar cookie group it was 20.11538462, and the control group's average after score was identical to its before score at 18.2333. As shown, the people who ate the no sugar cookie had a decrease of score, meaning their self-worth dropped, by 0.708333 points. The people who ate the sugar cookies' score dropped by about 0.11538461 points. And, the people who ate nothing actually had the exact same score, to the eighth decimal point. This suggests correlation between eating the food and the test scores.

CONCLUSION

The purpose of this study was to determine if refined sugar had an immediate effect on self-worth. The hypothesis for this study predicted that the teenagers who consumed the cookie containing sugar would score lower on the Rosenberg Self Esteem Scale when tested an hour after eating the cookie.

The results displayed only a minor negative impact in the sugar cookie and sugarless cookie groups, while the control group was not impacted at all, according to the averages. This brings to a possible conclusion that there is a strong cause and effect correlation between sugar and levels of self-worth, in teenagers. However, contrary to the hypothesis, the outcome showed that both the sugar *and* sugarless cookies impacted levels of self-worth. Based on post-study research and the multitude of other studies that suggest the negative impact of excessive consumption of refined carbohydrates, flours' high glycemic index of 71 versus sugars' GI of 68 [13], may have contributed to the more significant effect of the cookie on self-esteem. Tests with a larger sample size and a low GI grain to substitute the flour should be made to more finely pinpoint the effects of refined sugar.

Implications

Realizing that refined carbohydrates, including but not limited to refined white sugar, are directly related to negative feelings of self-worth in the short term can raise public awareness and desire to limit sugar intake, prompt doctors to recommend lowering intake as a supporting treatment for depression (which is strongly related with self-esteem [10], and could eventually lead governments to limit the amount of refined sugar permitted in food items which would greatly improve the health of the citizens. Much of the general public would consider said information to be important, since refined carbohydrates and refined sugar not only impacts our bodily functions, but also how we feel, which effects our work life, relationships, and perspectives of the world.

Future Work

This research suggests that refined sugar may have an impact on self-worth, but it is also likely that white flour has an immediate effect on short term changes in self-worth. It would be interesting to see if the results were the same with a larger sample size, leaving less chance for misled answers.

It would be interesting to repeat this experiment using ingredients that have a low GI index to define whether it was the refined white sugar alone, or any sugar at all including refined carbohydrates, that affected the self-worth of participants. Also, it would be interesting to see if artificial sweeteners had the same effect, which would help to understand if the rise in blood sugar was the only factor that decreased self-image.

In addition, experimenting with the amount of sugar consumed or the time interval between consuming the cookie and the end RSES assessment, may result in interesting differences in the data, and this information about the short-term effects of sugar could lead to a larger understanding of mental health. I also wonder if the time interval between the eating of the cookie and the end test would have caused more deviance in the before and after results, if it was longer and allowed for more digestion.

ACKNOWLEDGMENTS

The author wishes to thank their Mentor, Mrs. Lutovsky for her mentorship in proofreading and commenting on this paper, and Nathan Kartchner for encouragement through the difficulties of this project, as well as Mr. Napoli, Ms. Otani, Mrs. Wagner, and Ms. Appleby, for allowing me to experiment with students during their class time.

REFERENCES

[1] Avena, N. M., Rada, P., & Hoebel, B. G. (2007). Evidence for sugar addiction: behavioral and neurochemical effects of intermittent, excessive sugar intake. Neuroscience and biobehavioral reviews, 32(1), 20-39.

[2] Kenner, Robert, et al. (2008) Food, Inc.

[3] "Added Sugars." *Www.heart.org*, www.heart.org/en/healthy-living/healthy-eating/eat-smart/sugar/added-sugars.

[4] "Self-Esteem: The Nutrition Connection." *American Nutrition Association*, americannutritionassociation.org/newsletter/self-esteem-nutrition-connection.

[5] Johnson, R. J., Segal, M. S., Sautin, Y., Nakagawa, T., Feig, D. I., Kang, D. H., . . . Sánchez-Lozada, L. G. (2007, October). Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17921363

[6] The Immediate, Short And Long-Term Effects Of Sugar. (n.d.). Retrieved from https://www.thehealthyemployee.co.uk/news/the-immediate-short-and-long-term-effects-of-sugar

[7] Shaheen E Lakhan, & Annette Kirchgessner. (2013, August 08). The emerging role of dietary fructose in obesity and cognitive decline. Retrieved from https://nutritionj.biomedcentral.com/articles/10.1186/1475-2891-12-114

[8] Macht, M., & Mueller, J. (2007, November). Immediate effects of chocolate on experimentally induced mood states. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17597253

[9] Albrecht, Ashley. "The Effects of Self-Esteem and Stress on Eating Behaviours in Females." *Scholarship@Western*, ir.lib.uwo.ca/hucjlm/vol52/iss1/1/.

[10] Sowislo, Julia Friederike, and Ulrich Orth. " Does Low Self-Esteem Predict Depression and Anxiety? A Meta-Analysis of Longitudinal Studies." *Psychological Bulletin*, U.S. National Library of Medicine, Jan. 2013, www.ncbi.nlm.nih.gov/pubmed/22730921.

[11] (n.d.). Retrieved from https://openpsychometrics.org/tests/RSE.php

[12] Saunders, J., and Shannon Dailey. "The Best Rolled Sugar Cookies Recipe." *Allrecipes*, 2 Nov. 2006, www.allrecipes.com/recipe/10402/the-best-rolled-sugar-cookies/.

[13] Tremblay, S. (2018, November 19). Healthy Facts About Whole-Wheat Flour Vs. White.

Social Media and its effects on the Foreign Language Curriculum

Justine Kum
Dulwich College Seoul Banpo 2(i)-dong, Seocho-gu, Seoul, South Korea
Email: justinekum1004@gmail.com

***Abstract* – Learning a new language requires time, patience, and persistence. It is certainly not easy for many students. As Computer-Assisted Language Learning starts developing overtime through backgrounds such as social media, both benefits and problems arise surrounding the question on whether applying social media in an education curriculum would be useful. Overall, although useful, due to the inaccuracies portrayed in 'textese', social media seems to be more powerful for intermediate and advanced learners who want better cultural understanding rather than for teaching beginners who have to know standard 'rules' of a new language in a more formal setting.**

Key Words – Computer-assisted language learning, foreign language, social media, textese

INTRODUCTION

With 51.2% of the World's population using the internet in by the end of 2018 [1], the rising significance of technological developments guides people to greater awareness on the diverse cultures present in our world today. In particular, social media users are now able to observe how people with different backgrounds live in their own community and shape their identities based on their culture, making our world more interconnected [2]-[3]. Similarly, learning a foreign language is a step for people to acquire profound, cultural knowledge as it allows people from different cultures to communicate effectively and understand the current worldwide, diverse perspectives [4].

In terms of education, as conversations in the foreign language depend time-to-time based on the topic being discussed, learning a foreign language can involve a different learning style for students compared to standard curricula, which are prescriptive to teach students a set of knowledge. Thus, a curriculum for foreign languages can be a specific example where social media can actually be implemented in education, as it effectively illustrates the diverse spectrum of cultures and their uniqueness, including foreign languages. Furthermore, it seems that the connection between social media and language acquisition is only getting stronger as the presence of social media becomes more dominant and essential in modern life, especially when regarding education of young adults [5]-[6].

I. Purpose Statement

By understanding the possible relationship between social media and learning a foreign language, the purpose of the research project is to investigate on the effects of social media when applied to a foreign language curriculum. For a detailed discussion, I will explore two research questions related to the purpose. What's more, I see an extension of my culturally and linguistically diverse home. They are:

1. How might models for implementing social media in foreign language curricula impact learning?

2. How does modern language used in social media affect foreign language learning ?

II. Definition of Terms

There seems to be a common perspective that the only social media platform that exists are social networking systems such as Facebook, or Twitter, ignoring other areas of social media and their possible influences on various investigations associated with social media [7]. This is due to the fact that there is no clear definition of 'social media.' Combining its general features, social media will be referred to in this paper as a communication method for user interaction, and connection share their knowledge with other users, creating a central interactive network around the world.

In this study, the term 'modern language,' will be defined as a computer-mediated language that portrays the regular communication between users of social media regardless of its formality. For example, abbreviations such as LOL (laugh-out-loud) and ASAP (as soon as possible) are some phrases that may often be mentioned when referring to "modern languages." As social media is a platform which functions by computer-mediated communication, 'modern language' is a particular feature of the language in social media which needs to be thoroughly considered.

III. Delimitations of the Study

For the purpose of this paper, I will include only those research studies that work under the assumption that all schools, students, and teachers have access to various technologies, such as computers, internet and social media sources. However, studies such as [8], regarding the benefits and problems a developing country may face when using

social media in academic platforms, indicate how one crucial issue in these countries is the "the lack of adequate technical infrastructures" (p. 381). There have been multiple reports stating that developing countries, such as Zambia, Nigeria, and others in Southeast Asia, do not have sufficient knowledge on the benefits of social media in order for to use it in academic libraries. Even if some people are aware of the advantages provided by social media, its inaccessibility in developing countries acts as a barrier for its application in academics. On the other hand, some countries, like China, do not have access to certain social media platforms due to political issues [9]. However, these countries will be disregarded in this study completely as the majority of the world's population has access to the internet and social media (51.2% [1]), making the author focus more on the actual effects of social media rather than its accessibility to certain populations and further evaluation on this aspect.

Also, although languages are specific for each country and so, different curricular standards can exist depending on the language being taught, this study aims to unify language learning to one whole 'foreign language' curriculum as the process of learning a foreign language itself is very much the same for all languages. Thus, when talking about 'foreign languages,' the paper refers to all languages that an individual cannot speak fluently, rather than specifying on one particular language since one's native and foreign language depends on their personal background which is not considered in the research analysis. This also means that the paper assumes that the suggestions being illustrated are universally applicable, indicating that social media use is identical for learning all languages around the world. The assumption is suitable for the context of the paper as the author aims to focus more on the impact of social media as an educational source overall, rather than considering the details and depending situations of its impact. However, these assumptions are certainly the core topics of future potential studies the author aims to investigate on.

RESEARCH QUESTION #1 - HOW MIGHT SOCIAL MEDIA CHANGE FOREIGN LANGUAGE CURRICULA FOR UNIVERSITY STUDENTS?

Typical when learning a foreign language, students start feeling anxious as they think that they are incapable, especially since it is a time-consuming subject which requires thorough practice until mastered. This demotivates students to continue their language education [10]. Therefore, research on how to teach students foreign languages effectively so that students remain encouraged, motivated, and empowered to continue learning would be a beneficial study for foreign language educators. In turn, social media is a potential platform which can be used for this function, especially as modern students regularly use social media in their daily life. In the following section, I will consider the different features of social media to identify how they might benefit students in the language learning process.

I. Computer-Assisted Language Learning (CALL)

Studies in the involvement of technology in foreign language education, including ones by [5], [6], [11] and [12], all refer to the term 'Computer Assisted Language Learning' (CALL), the use of technology in foreign-language learning through media created for the specific purpose of educating its users. [13] researched on the incorporation of CALL in a typical classroom by having students, who are beginners of a foreign language, take a placement test to determine their language ability, and then use a software with various technological features that could help develop their skills of speaking, listening, writing, and reading. Some of the features included buttons such as the ""Exit" button … the "Microphone" button … and the "Headphone" button…. By recording the abundance of the students using each feature, the researchers could identify that the variables of CALL were effective in helping the participants answer more questions overall. Therefore, the results of [13] effectively supports the claim that CALL can be used as an advantage for student's education.

II. Social Media in Classroom Settings:

Although it has been identified that CALL can be a benefit to one's education of a foreign language, previous studies have also demonstrated its limits, most of which are based on students' motivation to actually use the specific language-learning platforms over more familiar and comfortable ones they use regularly.

For example, [14] conducted a survey "to examine students' perceptions and attitudes about using social media for learning foreign languages" (p. 65) and found that, out of the 720 university student questioned, around 570 (81%) use social media as a method of learning a foreign language. However, among the different options of platforms that social media offers, around 67% and 61% of the students would rather use social networks and media sharing sites respectively, whereas none of the students used language-learning specified technologies to develop their skills. This indicates how students use social media platforms that they are most familiar with, and use regularly outside the realm of education, rather than using those specified for curricular language instruction purposes. Teachers are advised to use social media platforms such as Facebook and YouTube during language education in order to motivate students. These are the most commonly used social media sites in the world, by having 85% and 79% of Internet users respectively [15], which allows more students to contribute to their learning.

III. Benefits of Social Media in a Foreign Language Curricula

[14]'s investigation of social media's impact on foreign language learning found that the foremost advantage of employing social media in the curriculum is having

students able to experience to communicate directly with native speakers rather than having uninteresting and 'confined' conversations in a classroom with other peers, who may probably be learners without proficient accuracy themselves. Such opportunities illustrate how various resources in social media, starting from social networking sites which allow global communication to broadcasting sites that expose students to how the language is regularly spoken in real life settings, are useful methods for their education. This revelation not only assists students' learning for the language itself, but also develops their cultural understanding by directly communicating with people from other cultures about varying perspectives. This is significant in foreign language learning as there is an "assertion that foreign languages are not only useful but necessary for an understanding of other peoples and other cultures," according to [4].

Results from research studies imply that services provided by various social media platforms are helpful for students who are learning a foreign language by developing both various skills, such as speaking, listening, reading and writing, as well as character, such as enhancing their self-esteem or an awareness of different cultures around the world [5], [14]. This once again indicates that social media can be very effective in the foreign language curricula. A survey carried out by [10] suggests that a student learning a foreign language would likely be fearful of learning the language as the process causes anxiety and confusion. Luckily, from the review of social media features that impact Polish foreign language students, it has been identified that the particular skills and self-growth acquired by utilizing social media as a method of learning can reduce the fear of a student learning a foreign language, and instead help them become proficient in applying the language in their regular life by advancing their communication abilities [16]. Therefore, the various platforms from social media acquiring different skills for an individual student can act as a key platform in providing the opportunity to develop language skills in diverse ways [14].

IV. Problems of Social Media in a Foreign Language Curricula

Although it has clearly been shown that social media can be applied in foreign language curricula for enhancing the education process for students, problems may arise when utilizing social media in this setting. [17] for example, examined the results of the Start Online project held in 2010, which attempted to implement social media in a pre-sessional program for new students at Bucks New University in order to improve the transitioning process from high school to university students had to face before their actual academic year started. The findings showed that when social media platforms were used for their intended purpose of developing social skills and interacting with other students, they were used extremely effectively. However, when attempting to use media platforms for academic activities directly, there was a significant downfall in the number of participants who were willing to engage in the study and interested in the process.

Similarly, with around 85% of the 502 students acknowledging that social media affects their life in some way, only 15% of the students in [18]'s survey responded that they have used social media for educational purposes. Additionally, [19]'s survey, which specifically focused on the impact of Facebook as a learning tool for students, said that the most uncommon use of Facebook was for their education, and therefore concluded that using Facebook in classrooms was an unsupported stance in their school.

What can be inferred from the results of these studies is that although the familiarity of social media can benefit a student's learning, this feature can distract, or even discourage them by applying educational purposes to something that students prefer to use for enjoyment and entertainment. This analysis is associated most directly with foreign language learning students as [14] stated that "it has been observed that though there are social online communities, such as Livemocha and Busuu, which are dedicated to language learning, none of the respondents seems to be using them,". Thus, students are not directly motivated to learn just by including social media into their curriculum and forcing them to use its features. Furthermore, it is emphasized that social media is a platform which users use regularly for entertainment and relaxation during their free time rather than for purposes for attentive tasks such as education.

RESEARCH QUESTION #2 - HOW DOES MODERN LANGUAGE USED IN SOCIAL MEDIA AFFECT FOREIGN LANGUAGE LEARNING FOR UNIVERSITY STUDENTS?

From 2001 to 2018, the percentage of individuals using the internet has increased from less than 10% to 51.2% [1]. Thus, it can be inferred that the growing use of the internet has also caused social media users to also increase, making the evaluation of the modern language used in social media necessary in order to identify whether social media is an effective platform which benefits or blemishes the foreign language learning process for university student.

I. Electronically Mediated Communication.

Another term used to describe communication using technology is Electronically Mediated Communication (EMC). There are various features of EMC which have been discovered since its creation that influence future developments of the English language [2]. As social media is a common form of electronically mediated communication, some general examples of the platforms associated with EMC's include emails, blogs, text messaging and instant messaging. [20].

Electronically Mediated Communication can also cause misunderstandings when sharing information. Research by [21] describes how EMC supports both simultaneous and

non-simultaneous conversation that can cause issues where the messages shared by the users are not necessarily received until much later. Researchers indicate how the gap in the timeframe from when the sender shares information, and when the recipient uses the information for their own understanding, can cause the meaning of the actual knowledge shared in the conversation to change. [23] defines this circumstance as "disembodied-language," and it is particularly crucial in acknowledging this feature of EMC when used in education. Teachers have to be aware of how to prevent misinterpretations of the information received by a student during their learning.

[2] signifies three ways in which EMC can impact students' learning of language:

1. EMC has "created a whole new language (p.162);"
2. Language is debased when created by technology;
3. Traditional language is harmed due to EMC.

Each of these elements will be investigated in later in this section of the paper.

II. Errors in Textism

An unusually abundant number of ellipses in sample texts led [20] to investigate 22 female undergraduate students at a public university in the American Midwest by recording the text messages they had sent in a single day. It was eventually identified that "ellipses are sometimes used to indicate speech trailing off… for dramatic effect… or to separate sentences in lieu of a more standard period (p.60)." Additionally, the search on exclamation marks and emoticons, in the same samples, showed how they could be applied to diminish the impact of periods, which express a brusque tone, by "softening or adding emotion to messages" (p. 61). The conclusion derived from this finding is that as punctuation is used in textese for demonstrating particular features of speech rather than acquiring to their conventional use of ending sentences or statements, the way they are used is informal and grammatically inaccurate.

III. Language Inaccuracy

The negative relationship between the use of inaccurate language in social media, mostly in textism, and users' ability to use the language in a linguistically correct way, helps demonstrate how social media platforms impact students, particularly when they are learning the assisted language through this application. Referring to [20]'s conclusion that adolescents may be able to learn writing skills through an electronic medium, the results of the studies mentioned in this literature review signify how informal language in social media influences the linguistic abilities of student users by making them too familiar to examples where language is used incorrectly for the simplicity. Although this also affects regular social media users, it can be further implied that foreign users who are unfamiliar with the language used are most largely affected. Foreigners are not aware of whether the language is simplified and set in an informal setting for convenience, or if the text portrayed reflects the conventional rules for the specific language. The imprecise boundary of the two situations is detrimental for foreign language learners as incorrect knowledge is being shared with the students, resulting in an improper language education system. Furthermore, it is necessary for users who are learning languages through social media platforms, including textism, to have an awareness of the possible inaccuracies in the language they are communicating in, so that they can determine whether they will continue using the platform for cultural knowledge acquisition, or find a more suitable method for learning.

However, even if the actual language the students write may have errors, [24] shows that these mistakes are not made intentionally as they are able to detect when it is appropriate to use textism. The corresponding study resulted from two findings. The first study was based on a questionnaire that asked undergraduate students when it is suitable to use textism in various settings, such as when writing emails to a friend, or submitting a final paper to a university professor. The results indicated that students recognized textism as appropriate in 'informal' settings, particularly in social settings, compared to those in conventional environments and situations. In particular, it was viewed that it was completely inappropriate for students to be using textism in exams, leading to the study's second finding of discovering how textism interacted with formal written exams.

The second finding showed that only 0.02% of the words in the exams had textism. This means that although there were common mistakes of missing capital letters, or apostrophes in the writing to create suspicion of the impact of textism in the English language, its scantiness allowed researchers to view the effect as negligible. In particular, the environments in which the students took the exam were external factors that could possibly have affected the students' performance, allowing this paper to make an assumption that the mistakes made in the texts may be insignificant.

In contrast, the same students reflected the behavior of using textism in their emails: "It is clear that even students who used textism in their emails were quite able to avoid using textism in their formal written exams" (p. 802). Thus, the study supports the finding that there is a close relationship between the regular use of social media and a user's language acquisition since the errors made are not intentionally done for convenient communication, but are actual mistakes caused by flaws in the user's knowledge of the language.

FINDINGS & IMPLICATIONS

Through thorough inquiry on the research questions by reviewing various studies carried out by multiple

researchers, this study was able to identify that social media can contribute in education systems and most specifically in foreign language curriculum. Moreover, by understanding how it is involved in such processes, it was possible to further investigate on how its contribution can be both beneficial and harmful for a student.

Social media is a beneficial platform as students are most comfortable and familiar with it. The various applications of social media systems allowed participants in studies to interact with those across the whole world. This is significant for foreign language learners as they are able to obtain the first-hand experience of communicating with native speakers, those who are the most knowledgeable and competent in utilizing the specific language. Furthermore, this benefit of sharing knowledge by multiple platforms in social media is an advantage that is useful for contemporary students as they are already familiar with the skills required to use the modern technology, and communication methods that enhances their engagement when using the system.

The second research question transitions to indicatively focuses on the language users of social media utilize regularly called 'Textese'. A flaw in using social media for learning a new language is the fact that it is defective when portraying rules which conventional language adheres to. As social media is a communicative system to make people's lives much more efficient, people have found ways to simplify it's language so that even though there are terminological inexactitudes in the text, the content of the message can still be shared. The impreciseness in the grammar, spelling and punctuation that 'textese' accommodates can cause students to be attuned to error-full social language, which also causes them to reflect these errors when the language is used in a different setting. More significantly, they are unable to identify when the language is a form of 'textese,' and when it is the traditional language, as they do not have any knowledge on its conventional structure. This makes social media a platform which does indeed educate students, but not necessarily with the correct information. As it is derived from our native language and influences communication both online and offline, textese is, furthermore, a suitable example to portray the evolution of languages. By becoming a regular communication language rather than a particular figure of speech on the internet, it can even be referred to as its own language for social media users, rather than a "flaw". However, in terms of the paper's purpose in identifying the effects of teaching a specific foreign language, 'textese' is an unsuitable platform for accurate education.

The most basic, but important, factor to consider is whether the students are competent in using modern technology and social media to implement its use in the curriculum. The advantage of social media discovered from the study is its familiarity, since it is assumed that most students regularly use social media in their daily life. Otherwise, the feeling of incapability when using such modern systems compared to others may cause a lack in engagement from the students in future studies. For this reason, it is also important for teachers to be aware of their own capacity in using social media as when there is a technological issue, or a student needs assistance in using the system, the first person students would proceed to for guidance are their teachers, and so the teachers should be prepared and able to adapt to the situations they face for supporting their students.

Because of the inaccuracies in social media teaching students incorrect language structures, there is also a premise that social media is more effective for the education of foreign language learners who are already quite experienced with using the language. These students are those who are moving to an advanced curriculum where they also have to start considering the cultural aspects of language learning. For beginners, it is more advised that teachers follow the conventional teaching method of a foreign language by teaching the basic grammatical rules of the language first. After the students have more knowledge of the vocabulary and are able to understand the language in more depth, it is then when applying social media to the education process is useful and effective. Furthermore, this provides the path for further research to focus on the possible impacts social media has on advanced language learners, and whether there are any differences in the overall outcome of teaching effectiveness for understanding what circumstances social media is useful for and suitable to use.

CONCLUSION

The findings from this research paper showed that the progressive increase in the development of technology and social media have led to its application in various fields around the world, including education. Currently, through its involvement in educating users, social media can provide both benefits and problems, particularly for foreign language learning students. Through developing an understanding of the diverging impacts social media can have on one's knowledge, the study provides the opportunity for readers to reflect on their regular use of social media and more specifically, the 'textese' they are associated with when uploading photos, making comments, or even liking a specific post, as the language and cultural context of these features can directly impact other users and the acquisition of their knowledge.

While it may be difficult to immediately determine how to take advantage of social media so that its problems are not reflected in educating students, it is important to acknowledge that the contribution of technology and social media is only growing in society, and so, attempts to utilize these platforms in curriculum are necessary for long-term development in foreign language curricula.

ACKNOWLEDGMENTS

The author wishes to thank Professor Jody Kerchner for her help and guidance in conducting the experiments and proofreading this paper.

REFERENCES

[1] United Nations ITU. 24 July 2019, https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.

[2] Baron, Naomi S. *Always on: Language in an Online and Mobile World*. Oxford University Press, New York;Oxford;, 2008;2010;, doi:10.1093/acprof:oso/9780195313055.001.0001.

[3] Harvey, Kerric, and Sage Publications. *Encyclopedia of Social Media and Politics*. SAGE Publications, Thousand Oaks, 2014.

[4] Brown, Roger W., et al. "Developing Cultural Understanding through Foreign Language Study: A Report of the MLA Interdisciplinary Seminar in Language and Culture." *Pmla*, vol. 68, no. 5, 1953, pp. 1196-1218.

[5] Alshenqeeti, Hamza. "Technology in the Language Classroom: How Social Media is Changing the Way EFL is Taught." *Arab World English Journal*, vol. 4, no. 4, 2018, pp. 55-68.

[6] Reinhardt, Jonathon. "Social Media in Second and Foreign Language Teaching and Learning: Blogs, Wikis, and Social Networking." *Language Teaching*, vol. 52, no. 1, 2019, pp. 1-39.

[7] Carr, Caleb T., et al. "Social Media: Defining, Developing, and Divining." *Atlantic Journal of Communication*, vol. 23, no. 1, 2015, pp. 46-65.

[8] Magoi, Joshua S., MK Y. I. Aspura, and A. Abrizah. "Social Media Engagement in Developing Countries: Boon Or Bane for Academic Libraries?" *Information Development*, vol. 35, no. 3, 2019, pp. 374-387.

[9] Huang, Paul. "Censorship in China Turns Social Media into Tool of Repression." *The Epoch Times*, 2018, pp. A9.

[10] Horwitz, Elaine K., et al. “Foreign Language Classroom Anxiety.” *The Modern Language Journal*, vol. 70, no. 2, 1986, pp. 125–132. *JSTOR*, www.jstor.org/stable/327317.

[11] Kessler, Greg. "Technology and the Future of Language Teaching." *Foreign Language Annals*, vol. 51, no. 1, 2018, pp. 205-218.

[12] K., Ganesh M., and S. K. Pushpalatha. "Defining the Roles of the Teacher and the Students in a Computer Assisted Language Learning Environment." *Language in India*, vol. 18, no. 10, 2018, pp. 371.

[13] Pinner, Richard S. "Teachers’ Attitudes to and Motivations for using CALL in and Around the Language Classroom." *Procedia - Social and Behavioral Sciences*, vol. 34, 2012, pp. 188-192.

[14] Faizi, Rdouan, Abdellatif El Afia, and Raddouane Chiheb. "Social Media: An Optimal Virtual Environment for Learning Foreign Languages." *International Journal of Emerging Technologies in Learning (iJET)*, vol. 9, no. 5, 2014, pp. 64-66.

[15] Ahmad, Waseem, and Rashid Ali. "Social Account Matching in Online Social Media using Cross-Linked Posts." *Procedia Computer Science*, vol. 152, 2019, pp. 222-229.

[16] Prizel-Kania, Adriana. "Surfing the Global Network! how to Incorporate Ict and Social Media in Teaching and Learning Foreign Languages." *Procedia - Social and Behavioral Sciences*, vol. 174, 2015, pp. 2777-2782.

[17] Knight, John and Rebecca Rochon. "Starting Online: Exploring the use of a Social Networking Site to Facilitate Transition into Higher Education." *Electronic Journal of E-Learning*, 2012, 10 (3): 257.

[18] Saha, S., Guha, A. Impact of Social Media Use of University Students, *International Journal of Statistics and Applications*, Vol. 9 No. 1, 2019, pp. 36-43. doi: 10.5923/j.statistics.20190901.05.

[19] Fewkes, Aaron M., and Mike McCabe. "Facebook: Learning Tool Or Distraction?" *Journal of Digital Learning in Teacher Education*, vol. 28, no. 3, 2012, pp. 92-98. d

[20] Baron, Naomi S., and Rich Ling. "Necessary Smileys & Useless Periods." *Visible Language*, vol. 45, no. 1/2, 2011, pp. 45.

[21] Chun, Dorothy, Bryan Smith, and Richard Kern. "Technology in Language use, Language Teaching, and Language Learning." *The Modern Language Journal*, vol. 100, no. 1, 2016, pp. 64-80.

[22] Pierce, Tamyra. "Social Anxiety and Technology: Face-to-Face Communication Versus Technological Communication among Teens." *Computers in Human Behavior*, vol. 25, no. 6, 2009, pp. 1367-1372.

[23] Clark, Herbert, H. “How do real people communicate with virtual partners? Proceedings of 1999 AAAI Fall Symposium.”, *Psychological Models of Communication in Collaborative Systems (pp. 43–47)*.1999, MA: AAAI

[24] Grace, Abbie, et al. "Undergraduates’ Attitudes to Text Messaging Language use and Intrusions of

Textisms into Formal Writing." *New Media & Society*, vol. 17, no. 5, 2015, pp. 792-809.

Equalizing Energy In The Food-Energy-Water Security Nexus: An Integrated Global Health Approach Using Reconstructed Assessment And Bioenergy Experimentation

Pranav Kanmadikar
DuPont Manual High School, Louisville, KY, USA
Email: pkanmadikar@gmail.com

Abstract **- The food-energy-water (FEW) security nexus is a crucial indicator of global health through its consideration of many multifaceted issues and recognition of inherent relationships between the usage of different resources. This research highlighted the FEW nexus through both qualitative assessment and quantitative laboratory approaches. Specifically, it was concerned with: (1) micro-level nonprofits' inaccessibility of nexus efficiency assessments; (2) how energy security is currently marginalized in nexus research and public advocacy; and (3) the approaches which can be undertaken to alleviate those issues. A qualitative approach was taken to modify a current nexus intervention assessment and to administer it to micro-level nonprofits, and a case study approach evaluating two water security organizations determined if energy security is marginalized within their current interventional efforts. Energy as a nexus issue was shown to not be adequately accounted for academically,publicly, and organizationally compared to food and water. As a solution, research was conducted specifically on biofuels, which sustainably integrate all three aspects of the nexus to produce two simultaneous energy sources—biogas and hydrogen—from municipal wastewater.**

Key Words–Food-energy-water nexus, Sustainability, Bioenergy, Intervention, Wastewater

INTRODUCTION

The association of global health insecurity with poverty is extremely prominent in issues of basic resource allocation of essential commodities. Social determinants of health generally include rates of safe food, energy, and water consumptionand whether those usage levels are sustainable for future years. Therefore, quantifying and utilizing the concepts of global resource *security*, or holistic access, safety, and sustainability, for food, energy, and water systems across the world is necessary to both alleviate health management generally as well as the social determinants which act as predecesso

massive dilemma affecting global health stakeholders lies in the varied ability to balance the sometimes-competing aims of resource security: the need for greater access to these commodities, coupled with expanding safety and health quality necessities.

One of the ways in which this dichotomy is addressed is through *food insecurity* classification. Unlike many other nutritional indicators, food insecurity rates have strikingly inclusive criteria thatencompass any person lacking access at any time to the necessary supply of food, including fresh fruits and vegetables that are crucial to a healthy diet. Perhaps this accounts for its prevalence, which the World Health Organization estimates to be at "[n]early a billion people," disproportionately within sub-Saharan Africa and the rest of the Global South [1]. Although there is sufficient nutritive capacity in the world today to feed all individuals, inequality in food distribution and production technologies leads to many areas having below-average consumption rates. The three criteria for food security, as set by the World Health Organization, are food availability, access, as well as utilization and diet, accounting for all possible competitions of interest for stakeholders in this issue and ensuring that interventions make holistic attempts to promote sustainability rather than just the discrete maximization of one indicator and neglect of other connected issues [1, 2].

Another resource concern that also integrates multiple separate factors is *water insecurity*. Like food security, water security measures if there is sufficient access towards, but also the safety and purity of, drinking water. Despite the prevalence of irrigation and fertile terrain in most of South and Southeast Asia, for example, those regions are among the most water insecure due to arsenic groundwater contamination as well as a lack of desalination mechanisms [3, 4, 5]. Besides this, the United Nations (UN) also includes water-borne pollution and disasters when defining water insecurity [6]. The multiple indices result in greater difficulty in terms of quantifying water insecurity, unlike food insecurity [7]. However, its scope is significant, with about 3.4 billion people either lacking direct access or the infrastructural and sanitation components necessary forsafe water consumption. This perpetuates many global health

issues, such as the spread of infectious or diarrheal diseases as well as higher infant mortality rates [8].

Moreover, one additional resource often neglected by global health pedagogies but having implications on both food and water sustainability is energy. Similar to its counterparts, energy security entails key indicators such as adequacy, affordability, and reliability. Additionally, energy type is even more significant regarding energy security because, despite their accessibility, some energies such as traditional biofuels have resulted in the deaths of millions. In those cases, household air pollution has often a prime contributing factor [9, 10]. Overall, the lack of energy access has ramifications on both of the other resources mentioned due to the energy needed for food production and water purification technologies.

The sum of all of the interlinkages between the largely overlapping aims of food, energy, and water security is known as the food-energy-water security (FEW) nexus. Nexus approaches specifically analyze the synergies and trade-offsbetween each combination of these resources. These mutual interconnections are crucial to address because of the fact that an innovation could indeed be advantageous for one component of the nexus, but without evaluating its implications on the other two, there is no way of knowing whether it is holistically advantageous [11]. One example of a synergy is the increase in both hydropower production (energy) and water management accompanying “large-scale water infrastructure projects,” while a trade-off can manifest as overreliance on imported food in response to water unavailability [12, 13]. The FEW nexus was chosen as the model for this research because of its broader implications on the fields of environmental epidemiology (e.g., through studying the spread of water-borne diseases) and health management. Therefore, the nexus can act as a measure of transdisciplinary integration into global health, and, by extension, utilize more methods to measure program effects on the central, as well as other, indicators.

However, there are barriers to the wider implementation of nexus interventionsbecause of a lack of assessment of global health nonprofits specifically. Although many indices have been published to evaluate integration within the nexus for general interventions, the tool created by the Food and Agriculture Organization (FAO) that uses “resource use efficiency indicators” (statistics linking two nexus components) is especially novel. (This is not for addressing environmental sustainability of the initiatives, but rather their efficiency with regard to all nexus issues) [14]. Despite its quantifiability, the FAO intervention assessment only addresses initiatives with measurable macro-level percentage changes and thus is not applicable to many nonprofits that perform impactful work but cannot numerically measure their efficiency with regard to the nexus. Additionally, suggested nexus indicators found in the comprehensive assessment working paper are largely agro-centric, scaling the agriculture-energy nexus to the context of the FEW nexus and focusing on issues such as power irrigation, desalination, etc. [13]. Therefore, the first part of this research entailed: (1) de-quantification and scaling of the FEW intervention assessments for microeconomic nonprofits and (2) addition of two new intervention categories and conceptualization of corresponding performance indicators for two additional “intervention types.” It was hypothesized that the above-mentioned changes would allow for effective qualitative assessment of organizations.

Within academia, a broader problem concerning the nexus is the water-centric nature of significant FEW nexus research. Greater than 50% of research showed preferences for particular nexus components over others and one-fifth of the total showed a preference for the water sector specifically. This is in contrast to the mere 8% and 7% for the energy and food sectors, respectively [14]. A likely contributor to this is the fact that, historically speaking, the FEW nexus originated from the water-centric “integrated water resources management” approach and thus may still retain some of the same structuring methodologies and mechanisms. The nexus, though, is contingent upon providing food, energy, and water equal focus while evaluating interactions between them [14]. Although specific sectoral focus is occasionally beneficial, this should be distributed equally among all nexus components rather than causing a large disparity in favor of water research. Equal distribution is imperative because of the fact that advancements in one area can have trade-offs on the other and technological advancements in water security, for example, should be evaluated alongside energetic and food-related impacts to ensure optimal solutions prior to advocacy. Thus, the second part of this research tested the issues as follows: (1) if the water nonprofit case studies mentioned previously show a neglect of the energy sector and (2) if public opinion prioritizes water security as well and reflects lack of awareness on energy. It was hypothesized that both approaches would show lessened prioritization of energy needs compared to other nexus issues.

To combat the issue of energy deprioritization within the nexus, bioenergy research was chosen as a method of advancing the sector while integrating it into issues of food and water sustainability. Bioenergy, as a form of renewable energy, can harness the agricultural and water residues of biomass to generate sustainable energy. One specific framework through which this can be performed is through a microbial electrolytic cell (MEC), in which biogas (a combination of methane and carbon dioxide) is normally generated as a source of energy because of bacterial fermentation. However, after collection of municipal wastewater and subsequent addition of an electrolyte, MECs can add an entry point for electrons in the system and produce hydrogen—a far more efficient and energetically dense source—as well. This is because, when bacteria are placed in an anaerobic MEC, they cannot transfer their electrons obtained from digestion to

oxygen and need to find another source [15]. Through incorporating residual agricultural and hydrological components in municipal wastewater, this MEC research can provide an example of a nexus integration technique that efficiently produces energy. Therefore, the effect of the electrolyte sodium acetate ($C_2H_3NaO_2$) was tested on energy outputs and it was hypothesized that there would be a positive trend in terms of simultaneous gas and hydrogen production [16].

METHODS

I. Nexus Assessment Modification and Distribution

Guidelines from the FAO working paper were followed, with the omission of quantitative (percentage) questions [13]. All but one of the given performance indicators were used in addition to 16 new ones (partly modified or newly conceptualized). This held true for all agricultural or bioenergy-related intervention categories. Additionally, two new categories were created for this assessment: Food distribution and Water purification and management, for which a further 19 new indicators were included. For each indicator, weights (1-3) were assigned based on subjects' own priority while the indicator measure was classified as positively, negatively, or not affected by the organization. Data analysis guidelines were followed to determine "score" for each nexus category, which included scaling of weighted averages to a 1-5 range [17].

II. Energy Opinion Survey

The surveywas administered to a select sample (n = 86) with questions relating to demographics and environmental opinions as well as awareness of biofuels and renewable` energy.

III. Bioenergy Experimentation

In a beaker solution, 470 mL of water, 2 g of sodium acetate, sodium bicarbonate, monosodium phosphate, and disodium phosphate were added, in addition to varying amounts of acid or alkali depending on the initial pH of the solution. Nitrogen purging was completed for 5 minutes to create an anaerobic environment, which was followed by application of electricity (1 V). The data analysis tool (pulse-flow respirometer) was disconnected after 50 hours, after which biohydrogen and biogas data were obtained and tested.Basic chemical lab safety training was required to complete this project. Project quality was carefully ensured through maintenance of all control variables.

RESULTS

I. Nexus Assessment Modification and Distribution

Of the two water security nonprofits shown in Figures 1 and 2, both were able to successfully complete a qualitative version of the nexus assessment. Scaled scores of each nexus indicator were calculated and graphed for food, water, and energy. Energy efficiency is shown to be marginalized in both cases, with values of 3.5 and 3 indicating no significant sustainable effect for that nexus component.

FIGURE 1
QUALITATIVE NEXUS ASSESSMENT SCORE OF WATER NONPROFIT #1

FIGURE 2
QUALITATIVE NEXUS ASSESSMENT SCORE OF WATER NONPROFIT #2

II. Energy Opinion Survey

Demographics
The majority of the surveyed subjects were either in high school (63.2%, 55) or university-equivalent (24.1%, 21), and all were residents of the United States.

Ecopolitical views
As shown in Table 1, a plurality (44.8%, 39) of subjects considered themselves liberal on issues of environmental policy.

TABLE 1
POLITICAL VIEWS REGARDING ENVIRONMENTAL ISSUES

Political views regarding environment	Number	Percentage (%)
Liberal	39	44.8
Conservative	8	9.2

Moderate	21	24.1
Prefer not to answer	19	21.8

Sustainability priorities

As graphed in Figure 3, respondents were significantly more likely to prioritize energy security least as a sustainability goal (19) than for either food (4) or water (3) security.

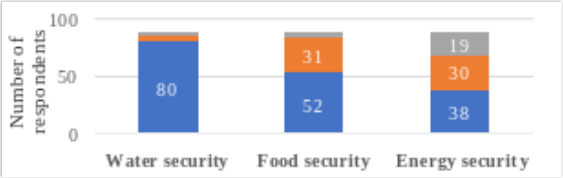


FIGURE 3
ENERGY DE-PRIORITIZATION AMONG SAMPLED SUBJECTS

Renewable energy awareness

Table 2 examines scholastic education on renewable energy. A 12.6-percentage-point difference was found between those who have not received in-depth renewable energy instruction (56.3%, 49) and those who have (43.7%, 38).

TABLE 2
VARIABLE SCHOLASTIC INSTRUCTION

In-depth renewable energy instruction (> 1 hour) at school	**Number**	**Percentage (%)**
Yes	38	43.7
No	49	56.3

Furthermore, awareness of biofuel renewable energy, in particular, was not necessarily correlated to school instruction. Table 3 demonstrates that a majority (59.8%, 52) were not educated in school about biofuels; however, a significant majority (70.1%, 61) are aware of their basic premises.

TABLE 3
AWARENESS OF BIOFUEL RENEWABLE ENERGY

Basic awareness of biofuel energy	**Number**	**Percentage (%)**
Yes	61	70.1
No	26	29.9

Despite a general lack of in-depth instruction in schools, most subjects (82.8%, 72) were aware of the fact that renewable energy, if proliferated, could alleviate climate change, Table 4 shows. Other data indicate that a vast majority (92%, 80) of respondents would have a positive reaction to more awareness and attention towards renewables in school.

TABLE 4
CLIMATE CHANGE MITIGATABILITY AWARENESS

Renewable energy can significantly mitigate climate change	**Number**	**Percentage (%)**
Yes	72	82.8
No	3	3.4
Not sure	12	13.8

III. Bioenergy Experimentation

As quantified in Table 5, all three sodium acetate concentrations were able to produce biogas and hydrogen simultaneously in MECs. As sodium acetate concentration increased from 2 to 6 g in otherwise similar MECs environments, biogas production decreased significantly (from 260.68 to 146.1 mL), while hydrogen production slightly grew (from 24.79 to 29.81 mL).

TABLE 5

SIMULTANEOUS BIOGAS AND HYDROGEN PRODUCTION FROM MECs TREATED WITH 2, 4, AND 6 g OF SODIUM ACETATE

Sodium Acetate (g)	**Biogas (mL)**	**Hydrogen (mL)**
2	260.68	24.79
4	220.76	29.78
6	146.1	29.81
Mean production	209.18	28.127

CONCLUSION

The nexus assessment administered to two nonprofits showed that the qualitative evaluation was effective and nexus component values were calculated for water, energy, food/land, labor, and costs. Additionally, these results added an organizational dimension to the problem of energy de-prioritization as both of the tested water security initiatives were shown to prioritize and significantly affect most other components, compounding the academic marginalization of energy that was most likely a result of the water-centric origin of a FEW precursor index. All research hypotheses were accepted based on the data.

In addition, the public poll demonstrated significantly more individuals placing energy as a sustainability issue of least importance. Despite this, there is awareness of renewable energy's climate change-related benefits and greater advocacy for it in schools is supported by a large majority.

Therefore, the resolution involved a nexus-based intervention attempting to maximize an existing bioenergetic mechanism. The MEC tested showed that both biogas and hydrogen could be produced in an electrolytic environment. Despite the significant difference in outputs between the two gases (as a result of hydrogen's greater density and safety), a combined solution will contribute greatly to energy security and

efficiency. In this case, greater sodium acetate optimizes total energy production and can contribute sustainably to non-traditional biomass.

Future research in this field may concentrate on a variety of issues, including the potential for the applicability of these nexus assessments to a broader range of both micro- and macro-level nonprofits (e.g., interventions for energy security, food security, or general sustainability). In addition, statistical modeling and quantitative conversion (in terms of measuring the extent of FEW impacts and consumption), if conducted in later research, will be able to further the credibility of the nonprofit assessment and, by extension, the nexus approach as a whole. Overall, it is imperative that such holistic lenses are utilized within the emerging academic field of sustainability prioritization in order to maximize the ability of nonprofits in the field to respond to scarcity and other global crises that are escalating well into the 21st century.

ACKNOWLEDGMENTS

I thank Paul M. Lewis and Bonny Lemma for their assistance with research procedures and writing of this paper as well as Dr. JagannadhSatyavolu and Dr. Samsudeen N. Mohamed for their research support of the bioenergy component of the project specifically.

REFERENCES

[1] Peace Corps. (n.d.). Global Issues: Food Security. Retrieved from https://www.peacecorps.gov/educators/resources/global-issues-food-security/

[2] Dhurandhar, E. J. (2016). The food-insecurity obesity paradox: A resource scarcity hypothesis. *Physiology & Behavior,162*, 88-92. doi:10.1016/j.physbeh.2016.04.025

[3] Abedin, M. A., Habiba, U., & Shaw, R. (2013). *Water insecurity: A social dilemma*.

[4] France-Presse, A. (2016, August 30). More than half of South Asia's groundwater too contaminated to use – study. Retrieved from https://www.theguardian.com/world/2016/aug/30/more-than-half-of-south-asias-groundwater-too-contaminated-to-use-study

[5] Ayers, J. C., &Goodbred, S. (2017). Arsenic Contamination in South and Southeast Asia. *Oxford Bibliographies Online Datasets*. doi:10.1093/obo/9780199363445-0087

[6] UN Water. (2013, May 8). What is Water Security? Infographic. Retrieved from http://www.unwater.org/publications/water-security-infographic/

[7] Boateng, G. O., Collins, S. M., Mbullo, P., et al. (2018). A novel household water insecurity scale: Procedures and psychometric analysis among postpartum women in western Kenya. *Plos One,13*(6). doi:10.1371/journal.pone.0198591

[8] Freshwater Project International. (n.d.). Global Water Crisis. Retrieved from https://www.freshwaterintl.org/global-water-crisis

[9] Yumkella, K. K. (2014, August 21). Need for Global Energy Security. Retrieved from https://worldpolicy.org/2014/08/21/need-for-global-energy-security/

[10] Kukreja, R. (2017, January 04). Causes and Effects of Indoor Air Pollution. Retrieved from https://www.conserve-energy-future.com/causes-and-effects-of-indoor-air-pollution.php

[11] Nielsen, T., Schhnemann, F., Mcnulty, E., et al. (2015). The Food-Energy-Water Security Nexus: Definitions, Policies, and Methods in an Application to Malawi and Mozambique. *SSRN Electronic Journal*. doi:10.2139/ssrn.2740663

[12] Food and Agriculture Organization of the United Nations. (n.d.). The Water-Energy-Food Nexus. Retrieved from http://www.fao.org/3/a-bl496e.pdf

[13] Flammini, A., Puri, M., Pluschke, L., et al. (2014). *Walking the nexus talk: Assessing the water-energy-food nexus in the context of the Sustainable Energy for All initiative*. Rome: Climate, Energy and Tenure Division (NRC), Food and Agriculture Organization of the United Nations.

[14] Albrecht, T. R., Crootof, A., & Scott, C. A. (2018). The Water-Energy-Food Nexus: A systematic review of methods for nexus assessment. *Environmental Research Letters,13*(4), 043002. doi:10.1088/1748-9326/aaa9c6

[15] Logan, B. E. (n.d.). Microbial Fuel Cells. Retrieved March 02, 2018, from https://www.engr.psu.edu/ce/enve/logan/bioenergy/research_mfc.htm

[16] AyushAyush 1912, Oscar LanziOscarLanzi 15.9k12648, et al. (2017, October 29). Is sodium acetate strong or a weak electrolyte? Retrieved from https://chemistry.stackexchange.com/questions/84955/is-sodium-acetate-strong-or-a-weak-electrolyte

[17] User650271user650271 1, Irritateirritate 5, Charles ClaytonCharles Clayton 7, et al. (14, March 2011). How to scale down a range of numbers with a known min and max value. Retrieved from https://stackoverflow.com/questions/5294955/how-to-scale-down-a-range-of-numbers-with-a-known-min-and-max-value

SUPPLEMENTALS

In order to conduct the project, the following indicators were utilized for water purification interventions or parts of interventions:Δ Water yield / energy required for water structure construction, Δ Energy consumed / water purification machines, Δ Fossil energy consumed / water purification machines, Δ Renewable energy consumed / water purification machines, Δ Land occupied / water treatment facility, Δ Treated water use / irrigation and agricultural purposes, Δ Total hours saved from extracting and carrying water / land, Δ n. of skilled jobs /

clean water plant installed, Δ Cost / unit of treated water for farmers, Δ Value of treated water / annual operating and capital cost, Δ Capital and cost expenditure for equipment / cost of workforce, Δ Energy required to process residues and toxins from contaminated water / disposal facility on land, and Δ Costs / disposal facilities for water contaminants.

In the case of Water Security Nonprofit #1, performance indicators were weighted and typified after the survey, as follows on Table 6:

TABLE 6

PERFORMANCE INDICATORS FOR WATER SECURITY NONPROFIT #1

Nexus Component	Performance Indicators (nexus component, chosen weight) (type of impact)
Water	Δ Water yield / energy required for water structure construction (W, 3) (+)
	Δ Treated water use / irrigation and agricultural purposes (W, 2) (+)
	Δ Value of treated water / annual operating and capital cost (W, 3) (+)
Energy	Δ Energy consumed / water purification machines (E, 1) (0)
	Δ Fossil energy consumed / water purification machines (E, 1) (0)
	Δ Renewable energy consumed / water purification machines (E, 1) (0)
	Δ Energy required to process residues and toxins from contaminated water / disposal facility on land (E, 3) (+)
Food/Land	Δ Land occupied / water treatment facility (F, 2) (+)
Labor	Δ Total hours saved from extracting and carrying water / land (L, 3) (+)
	Δ n. of skilled jobs / clean water plant installed (L, 2) (+)
Costs	Δ Cost / unit of treated water for farmers (C, 3) (+)
	Δ Capital and cost expenditure for equipment / cost of workforce (C, 3) (+)
	Δ Costs / disposal facilities for water contaminants (C, 1) (0)

Scaled scores were as follows on Table 7 for Water Security Nonprofit #1:

TABLE 7

SCALED SCORES FOR WATER SECURITY NONPROFIT #1

Nexus Component	Score
Water	4.8
Energy	3.5
Food/Land	4.3
Labor	4.7
Costs	4.6

In the case of Water Security Nonprofit #2, there were multiple intervention components besides water purification, but only the latter was tested for the sake of consistency. Categorization was as follows on Table 8 to determine raw scores:

TABLE 8

PERFORMANCE INDICATORS FOR WATER SECURITY NONPROFIT #2

Nexus Component	Performance Indicators (nexus component, chosen weight) (type of impact)
Water	Δ Water yield / energy required for water structure construction (W, 2) (+)
	Δ Treated water use / irrigation and agricultural purposes (W, 1) (0)
	Δ Value of treated water / annual operating and capital cost (W, 1) (0)
Energy	Δ Energy consumed / water purification machines (E, 2) (0)
	Δ Fossil energy consumed / water purification machines (E, 1) (0)
	Δ Renewable energy consumed / water purification machines (E, 1) (0)
	Δ Energy required to process residues and toxins from contaminated water / disposal facility on land (E, 1) (0)
Food/Land	Δ Land occupied / water treatment facility (F, 2) (+)
Labor	Δ Total hours saved from extracting and carrying water / land (L, 3) (+)
	Δ n. of skilled jobs / clean water plant installed (L, 3) (+)
Costs	Δ Cost / unit of treated water for farmers (C, 1) (0)
	Δ Capital and cost expenditure for equipment / cost of workforce (C, 1) (+)
	Δ Costs / disposal facilities for water contaminants (C, 1) (0)

Scaled scores were as follows on Table 9 for Water Security Nonprofit #2:

TABLE 9

SCALED SCORES FOR WATER SECURITY NONPROFIT #2

Nexus Component	Score
Water	3.4
Energy	3.0
Food/Land	4.3
Labor	5.0
Costs	3.2

Reduction in air pollution to build a sustainable global society

Winston Jim Lam
English Schools Foundation West Island School, Hong Kong, China
Email: 19lamw2@wis.edu.hk / winstonj.lam@gmail.com

***Abstract* - This research paper discusses the problems humanity faces with regard to air pollution and its severity through three perspectives/lenses: natural science, global public health and international policy. With risks and threats underscored by research in the natural sciences and public health, the effects of government policies cannot be ignored. This paper concludes with possible courses of action to deal with the issues with reference to modern society.**

Key Words – Air pollution, humanity, sustainability

INTRODUCTION

Research question – Having analysed the current global situation, in what ways can humanity reduce air pollution to build a sustainable society?

Humanity's rapid development and industrialisation since the first Industrial Revolution has made living substantially easier than it was just centuries ago. Nevertheless, a major issue emerges from such exponential growth: air pollution. According to the United Nations, a vast majority of 92% of the global population live in places where air pollution exceeds accepted levels [1]. If the current rate of air pollution production continues to grow, humanity is confronted with a grave danger accelerating its extinction. This paper discusses the issues currently faced by, the research and the possible courses of action regarding air pollution, with a framework based on three societal 'lenses': the natural sciences, global public health and international policy, with reference to the UNSDG goals. (Fig. 1)

FIGURE 1
UNSDG GOALS REFERENCED IN THIS PAPER
Source: "Sustainable Development Goals." United Nations, United Nations, www.un.org/sustainabledevelopment/sustainable-development-goals/.

LITERATURE REVIEW

The Oxford Dictionary defines air pollution as 'the presence in the air of any air pollutant that reduces air quality enough to threaten the health and welfare or people, plants, and animals.' This paper also considers sustainability as 'the ability [for humanity and its processes] to be maintained at a certain rate or level'.

According to the World Health Organisation (WHO), the underlying issue is that many developing countries continue using non-renewable fuels, while the developed world has been generally inactive in switching to renewable fuels, given that outdoor air pollution is from the 'inefficient combustion of fuels for transport, power generation and other human activities' [2] (Fig. 2). Scientific consensus is that that environmental pollution can only be prevented, not reversed. Environmental science, medical and international policy studies strongly support this consensus. Both particulate matter and oxides of nitrogen create immense problems with climate change affecting human and environmental health. It echoes the importance of pre-emptive anti-air pollution policy, such as adherence to global climate accords such as the 1997 Kyoto Protocol and the 2016 Paris Accord as international policy. However, neither of these have been universally ratified nor seen great success. This increases global sentiment on the dire nature of the air pollution problem, across academic disciplines and the public alike.

This paper, through the three lenses, compares and contrasts three countries in case-study format (Germany, China and the United States) to acquire a holistic and international understanding of air pollution.

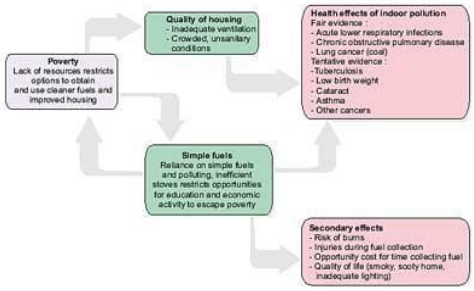


FIGURE 2
FLOW CHART OF THE CAUSES AND EFFECTS OF AIR POLLUTION [2]

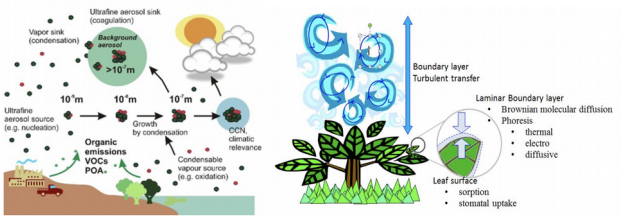
LENS 1: THE NATURAL SCIENCES

Scientific analysis of global air pollution and the pollutants within it gives insight into its effect on the Earth and the risks it poses to the human race. This section focuses on **two criteria pollutants**: particulate matter (PM) and oxides of nitrogen (NO_x).

Particulate matter, or atmospheric aerosol, is a mixture of microscopic particles that create the largest risk to humans and the environment among the many components of air pollution. It is comprised mostly of

sulphates, nitrates, ammonium, elemental and organic carbons, silicon and sodium ions. [3] Scientific literature on this pollutant, especially that of 2.5 microns termed $PM_{2.5}$, has increased dramatically in recent years. [4] $PM_{2.5}$ is produced by many sources, often of an anthropogenic nature including wood combustion, transportation, and factory emissions. [5] Its nature of amalgamating in the atmosphere and **nucleation scavenging** or **wet depositing** (Figs. 3 and 4) in precipitation forms a cycle of $PM_{2.5}$ which purports large concentration across the globe. Because of this cycle, with reference to studies in China, $PM_{2.5}$ is most hazardous during dry days, especially in winter. Without proper ventilation and precipitation, PM coagulates and stagnates low in the troposphere, forming a thick and harmful fog. [6] Moreover, $PM_{2.5}$ spreads over time due to natural trade winds, creating a global problem.

Apart from human health, $PM_{2.5}$ also poses great danger to the health of the environment and its ecosystems. It depletes soil nutrients and damages forests and crops. Scientists most strongly correlate this to **dry deposition** of $PM_{2.5}$ on plants. [4] Collectively, wet and dry deposition impede plant growth by coating upper leaf layers blocking light exposure and gas exchange. [7] $PM_{2.5}$ infiltrates all ecosystem levels: biochemical studies show that lichens (environmental health indicators) and detritivores in the soil are all affected because $PM_{2.5}$ components (such as MgO) alter pH. Noting that there is an about 10-year lag in data collection, ascertaining $PM_{2.5}$'s true effects scientifically is considerably difficult. Clearly, $PM_{2.5}$ poses severe risks, further complicating paths to achieving UNSDG Goals 11 and 13.



FIGURES 3 AND 4
SCHEMATICS SHOWING THE **AMALGAMATION** AND **DEPOSITION** OF $PM_{2.5}$ IN THE ATMOSPHERE AND ONTO PLANTS RESPECTIVELY [4]

Oxides of nitrogen (NO_x), including nitrogen dioxide (NO_2) and nitrous oxide (N_2O), are primary pollutants which are attributed to factory production. NO_2, in particular, is derived from petrol engines in vehicles. [8] NO_x, like $PM_{2.5}$, damages both health and the environment, causing the formation of **acid rain**, dense fog and nutrient pollution in bodies of water. NO_x also creates substantial problems with ozone; [9] while ozone levels naturally fluctuate due to oxygen-ozone dynamic equilibria in the stratosphere, the additional anthropogenic pressure of NO_2 damages the ozone layer in a cycle of chemical reactions (1), resulting in increased amount of UV light reaching Earth, causing human skin cancer and crop/marine life death. NO_x reactions (2) with volatile organic compounds (VOCs) form **tropospheric ozone**, a harmful pollutant that damages the health and foliage of trees and crops. Overall, the risks posed by NO_x worsen the environmental damage caused by air pollution.

$$NO_{2(g)} + \text{sunlight} \rightarrow NO_{(g)} + O_{(g)} \quad (1)$$
$$O_{2(g)} + O_{(g)} \rightleftharpoons O_{3(g)}$$
$$NO_{(g)} + O_{3(g)} \rightarrow NO_{2(g)} + O_{2(g)}$$
$$NO_{x(g)} + VOCs_{(g)} + \text{sunlight} \rightarrow O_{3(g)} \quad (2)$$
(+ other products)

LENS 2: GLOBAL PUBLIC HEALTH

According to the WHO, 7 million people die annually due to polluted air and 9 out of 10 people breathe it in per day. [10] Medical and public health analysis into air pollution reveals that it adversely affects human health, especially the elderly and children. Air pollution also contributes to the aetiology of many diseases, while some studies have even shown that excessive $PM_{2.5}$ exposure **reduces** human life expectancy by around 1-2 years. [11] This section focuses on the effects of air pollution on human health caused by particulate matter and NO_x.

Extensive medical studies into particulate matter have repeatedly suggested that it poses great risk to human health. Due to its microscopic size, particulate matter is able to penetrate deep into the lungs and bronchi, causing inflammation. They are also known carcinogens. [12] Smaller particles such as PM_1 (1 micron) can penetrate the air-blood barrier and spread throughout the body, leading to diseases of kidney, spleen and liver. PM's nature to amalgamate causes atherosclerosis in the blood stream, ultimately causing cardiovascular disease. (Fig. 5) [13] The situation in the three case-study countries is analysed below.

In Germany, extensive studies have been performed in paediatric and public health medicine to analyse the current air pollution situation. While Germany has worked to reduce $PM_{2.5}$ levels, life expectancy due to the current levels is still reduced by about 10.2 months. [14] Morgenstern et al. found that German children exposed to $PM_{2.5}$ generally saw an increased level of allergies and **atopic** diseases such as asthma and bronchitis, in that the likelihood of having asthma was 1.6 times higher than normal when being 50m from the main road. (Fig. 6) [14] Studies have also shown that there is an increased risk of about 1.36 times on diabetes prevalence upon exposure to an increase of 1 $\mu g/m^3$ of PM 2.5. It is clear that air pollution has more overarching health effects than previously anticipated. [15]

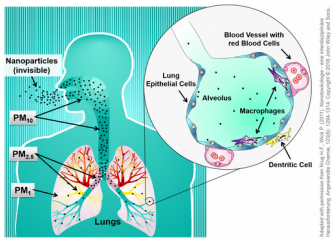


FIGURE 5
SCHEMATIC OF HOW PARTICULATE MATTER (OF ALL SIZES) ENTERS AND PENETRATES THE HUMAN RESPIRATORY SYSTEM [13]

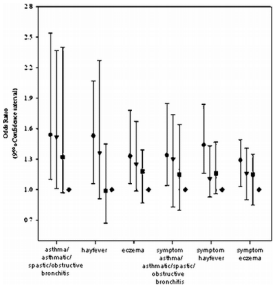


FIGURE 6

GRAPH WITH ERROR BARS OF DIFFERENT ATOPIC DISEASES AND ALLERGIES ATTRIBUTED BY $PM_{2.5}$ (FOUR BARS FROM LEFT TO RIGHT OF DIFFERENT DISTANCES FROM MAIN ROAD: 50M, 50-250M, 250-1000M AND >1000M RESPECTIVELY) IN GERMAN STUDY [14]

Findings in China, a country with significantly more severe air pollution than Germany, concurs with the German findings. W et al. found that exposure to $PM_{2.5}$ and PM_{10} was positively correlated to increased influenza-like symptoms in Jinan, especially for infants and adults. [16]. Studies conducted in China also have correlated diabetes and air pollutant exposure, with every 10 $\mu g/m^3$ increase in exposure of all sizes leading to increased mmol/L levels of blood glucose. [17] Furthermore, studies in Beijing have found strong positive correlations between $PM_{2.5}$ exposure and an increased likelihood of miscarriage in expectant mothers, affecting a new level of society. [18]

Similar findings are observed from studies in the United States. Studies in New England determined a **causal relationship** of $PM_{2.5}$ on mortality rate, with an IQR increase associated with a 4.04% death rate increase. [19] Studies in California found similar results (i.e. the same positive correlation) in adult cohorts. [20] Overall, a large body of scientific literature in the United States suggest high $PM_{2.5}$ exposure has risks of lung cancer and respiratory illness **and** diabetes, triangulating the aetiology of these diseases to $PM_{2.5}$.

NO_x, slightly different to PM, has known effects on human health due to its atmospheric chemistry and production method. NO_x is considered a primary air pollutant, as it also acts as a precursor to secondary air pollutants like O_3 and HNO_3. Ground-level ozone can severely damage lung tissue at the cellular level even with short-term exposure since it is a powerful oxidative agent [21] and in constant equilibrium with free radicals (1). NO_2 in particular causes great damage to human health. It does not only create harmful stratospheric ozone but is also a major component in photochemical smog. It is toxic and triggers cell damage along the respiratory tract. The photochemical smog mixed with tropospheric ozone both pose grave danger to human health. The diseases and damage caused by NO_2 and O_3 are analysed in the three case-study countries below.

Germany, a country with one of the strongest diesel automobile industries in Europe, suffers greatly from the afflictions from NO_2 and O_3. While studies have suggested an annual mean of 20 $\mu g/m^3$ is best for human health [22], Berlin and many other German cities consistently top such levels, even topping the 40 $\mu g/m^3$ EU limit. Studies published by the German Federal Environment Agency have found that even small levels of NO_2 are linked to diabetes [23], lung cancer, and premature birth. [24] Other studies correlated moderate levels of NO_x to increased risk of paediatric respiratory illnesses [25], and thus it affects all walks of life in Germany.

China, a country equally struggling with vehicular use, has extensively studied the effects of NO_2 and ozone on its people. China's industrialisation has dramatically increased ambient air NO_x due to increased coal combustion and vehicular use. [26]. Paediatric epidemiological studies in Chongqing have suggested NO_2 reduces foetal and child growth while increasing the risks of cancer. [27] In Shanghai, where NO_x is a major pollutant, studies have linked a 10 $\mu g/m^3$ increase in concentration with a 0.52% increase in hospital visits – scalable to 120,000 of Shanghai's residents. [28] Ozone has been predicted in China to be detrimental to public health (Fig. 7), concerning as 67.2% of the population live above the accepted WHO standards [29]. Linear projections predict 696,000 deaths due to ozone pollution, and that if ozone was reduced, nearly 80,000 deaths could be prevented. [30]

The United States' findings, much like those for $PM_{2.5}$, support those of Germany and China. Adams et al. [31] found that NO_2 exposure is not only linked to respiratory illnesses, but is also linked to olfactory dysfunction in older adults. Moreover, ozone is attributable to relatively high percentages of death from air pollution (10% in Los Angeles) [32], 'increases in daily mortality rates, increases in hospitalisations and exacerbations of respiratory illness'. [33]

Scientific literature from these three case-study countries can be said to give resounding evidence, for NO_x (and its secondary pollutants), air pollution in general to be considered severe and pressing risks to public health, impeding progress in achieving Goals 3 and 15.

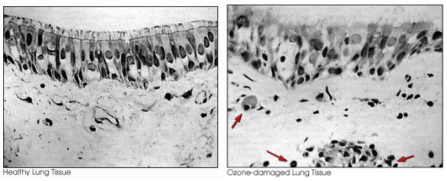


FIGURE 8

MICROGRAPHS OF HEALTHY (LEFT) AND TROPOSPHERIC OZONE-DAMAGED (RIGHT) LUNG TISSUE: EFFECTS OF MISSHAPEN CILIA AT TOP OF EPITHELIAL CELLS AND HIGHER PRESENCE OF NEUTROPHILS

Source: "The Ozone We Breathe." NASA Earth Observatory, NASA, earthobservatory.nasa.gov/features/OzoneWeBreathe/ozone_we_breathe2.php.

LENS 3: INTERNATIONAL POLICY

So far, this paper has presented evidence in both natural science and public health literature emphasising air pollution as a pressing issue. Focusing attention onto solving the problem though, international policy and united actions among countries on air pollution is really the only way to make change. Policy reviews are made of the three case-study countries below.

Germany is a relatively air pollution accord compliant country. It is a strong renewable energy advocate, and in 2016 acceded to the Klimaschutzplan, an

action plan for air pollution up to 2050 [34]. Chancellor Merkel, Germany's current leader, is a strong supporter of combatting air pollution: she has promoted the Energiewende and has been an advocate of the 2016 Paris Accord. German climate policy such as the Energiewende is generally successful. The Energiewende, a plan implemented to replace nuclear and fossil fuel power with renewable energy, was implemented in 2000. Its ultimate success is visible in that renewable energy now generates around 33% of Germany's total electricity and is nearly on par with coal in terms on total terawatt-hours produced. [35] Other German policy programmes, such as the Klimaschutzprogramm 2030 and the Klimaschutzplan 2050, implemented in 2014 and 2016 respectively, need to be analyse for longer before conclusions can be drawn, yet the former seems to be on track to reaching Germany's goal of greenhouse gas neutrality by 2050. [36] These policy successes show Germany is making steps forward in reaching UNSDG goals 3, 7 and 9, leading to real change.

By contrast, the United States, specifically under its current administration, challenges most air pollution consensuses. Its policies currently lean towards fossil fuel industry than public health.

President Trump's air pollution policy is a testament to his willingness to challenge scientific consensuses; an example of this is his decision in 2017 to pull the United States out of the Paris Accord, despite 70% of Americans wishing to stay. [37] His America First Energy Plan, aimed at bringing back coal mining jobs in rural America, has sparked controversy among environmental and public health scientists for its seeming ineptness and lack of understanding of the current crisis.

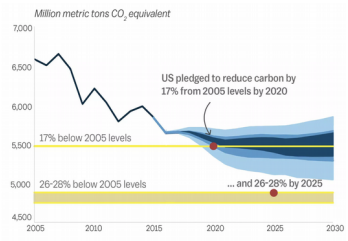


FIGURE 9
GRAPH OF US CARBON EMISSIONS OVER THE 21ST CENTURY (NOTE 2014-2019 FALL IN EMISSIONS MOST LIKELY DUE TO OBAMA-ERA POLICY AND THE PLATEAU THAT FOLLOWS UPON INITIATION OF TRUMP'S THE AMERICA FIRST ENERGY PLAN) [39]

Former President Obama's Clean Energy Plan was the first plan in United States history to put caps on carbon emissions from factories in America, responsible for nearly 40% of total US emissions. [38] It cut back funding for coal and increased funding for renewable energy. The plan was quite successful: the period of Obama's presidency (Fig. 9) saw an around 500 million tonne drop in CO_2 emissions. [39] Along with the Clean Water and Air Acts, this policy show the previous success the United States has had when it adhered to air pollution reduction policies. Nevertheless, Trump's current end to the 'War on Energy' has cast great uncertainty over America's place in solving the air pollution crisis. By cutting funding to the EPA by a third [21] coupled with fossil fuel production now back on the rise, it is unlikely that the United States will be able to reach its 2025 goal of sub-5000 million tonnes per annum. Overall, the case study of the United States underscores 'the importance of… long term research… as part of fact-based decision making in environmental policy.' [40]

China is the largest producer of air pollution in the world. It often suffers from pollution levels substantially above WHO levels, likely attributable to its inaction towards air pollution throughout its history in favour of economic progress. [41] The first major anti-air pollution law in China was not made until the late 1970s with the 1979 pollution levy, yet its implementation was impeded by many societal factors. [42] Coal consumption for household use and electricity generation remains high, using as much as the sum consumed by the rest of the world. [43] Initial Chinese policies, in comparison to current ones, failed to combat the air pollution problem.

Fortunately, China has moved forward in tightening controls on air pollution. The Air Pollution Prevention and Control Action Plan (APPCAP) set out by the State Government in 2013 was a milestone in Chinese air pollution reduction progress, and it was successful: in the 74 key cities studied, annual $PM_{2.5}$ concentrations fell by 33.3% and annual SO_2 concentrations fell by 54.1%. (Fig. 10) [44] President Xi announced China's continuation of its 'years-long battle against smog' for clear air and promoting a 'revolution in clean energy' in 2017. [45] China aims to cut coal usage by 58% by 2020, and is maintaining its commitments thus far. [46] The Chinese government, in contrast to the United States, stands firm against big industry to uphold sustainability and aims to reach UNSDG goals 3, 7, 11 and 13.

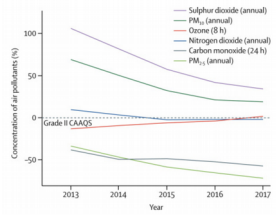


FIGURE 10
GRAPH OF AIR POLLUTANT LEVEL CHANGES IN CHINA AFTER THE IMPLEMENTATION OF THE STATE APPCAP POLICY

POSSIBLE COURSES OF ACTION

With reference to the research in this study, **four** main courses of action should be used in conjunction to tackle the air pollution crisis: (i) ensuring adherence to global climate accords, (ii) raising awareness on air pollution severity, (iii) quitting fossil fuels for renewable energy, and (iv) using innovative technologies. (Fig. 11)

Ensuring adherence to global climate accords is a difficult but crucial part of the solution. Countries such as Sweden, which recently signed a 'Climate Act' pledging even stricter regulations than it agreed to in Paris [47], and Germany, as discussed previously, are setting good examples for other countries on pollutant policy. It is

advisable that countries do more to reduce emissions to a level needed to contain climate change.

Secondly, fossil fuel combustion is central to the existing amount of air pollution, and thus its decrease and cessation would reduce emissions dramatically. Most countries still do not shift to renewable energy on economic considerations. However, countries like China are making progress in renewable energy development. Thus, even if the change from fossil fuels is more expensive, the shift should be considered a good investment towards the planet's sustainable future.

Thirdly, raising awareness about air pollution is an integral part in any course of action. Promotion among the public, at all ages, about the risks of air pollution not only helps people understand more about the dangers uncontrolled pollution can have, but also convinces them to rally behind those in power to take action and make actual change. The power of the people may truly be one of the most determining factors in defeating air pollution.

Lastly, **novel** and **innovative** solutions or ideas to help counter the air pollution crisis are emerging. For example, Climeworks, a Swiss company, is working towards building machines that can remove CO_2 from the atmosphere and turn it into stone for construction. Solutions such as [48] vertical forests in megacities, algae curtains and capturing air pollution to form ink cartridges (by Graviky Labs) are all pioneering technologies that, as the majority of them get cheaper, should be universally considered alongside the fundamental methods .

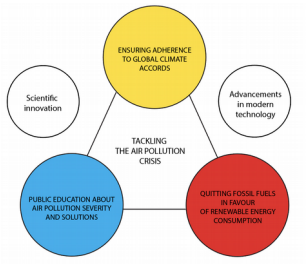


FIGURE 11
SCHEMATIC OF POSSIBLE COURSES OF ACTION AGAINST AIR POLLUTION

CONCLUSION

It is clear that air pollution is inexplicably and universally damaging; it requires stringent globally ratified control that **must start now**, for the damage and risks observed through with the three lenses studied in this paper (the natural sciences, global public health and international policy). The unwavering scientific evidence, the dire health risks and examples of successful combative policies serve to show UNSDGs 3, 7, and 11 (Fig. 1) can be achieved. It is strongly advised that both the innovative and general guidelines for air pollution solutions be heeded globally in hopes of actually solving the problem. Global concerns about air pollution have increased dramatically over the last decade. It would be destructive, if not tragic, if society recognises too late when ultimately air pollution grows too big. United must all facets of humanity stand in strong support of and as part of the air pollution solution.

ACKNOWLEDGMENTS

I would like to thank Ms Catherine Schofield for helping me write the primary submission of this paper as my iGCSE Global Perspectives coursework, and Dr. Rachel Yu, my iB Chemistry teacher, for providing this opportunity and giving constant support in HL Chemistry, which helped immensely in writing the natural science lens of this paper.

IN TEXT REFERENCES

[1] "6.76 Billion People Living with Excessive Air Pollution." *United Nations*, United Nations, www.un.org/sustainabledevelopment/blog/2016/09/vast-majority-of-world-6-76-billion-people-living-with-excessive-air-pollution-un-report/.

[2] "Air Pollution." *World Health Organization*, World Health Organization, www.who.int/ceh/risks/cehair/en/.

[3] Dominici, Francesca, et al. "Chemical Composition of Fine Particulate Matter and Life Expectancy." Epidemiology (Cambridge, Mass.), U.S. National Library of Medicine, July 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC4742572/.

[4] Fuzzi, S, et al. Particulate Matter, Air Quality and Climate: Lessons Learned And Future Needs. Atmospheric Chemistry and Physics, 24 July 2015, www.atmos-chem-phys.net/15/8217/2015/acp-15-8217-2015.pdf.

[5] "Particulate Matter (PM) Basics." *EPA*, Environmental Protection Agency, 14 Nov. 2018, www.epa.gov/pm-pollution/particulate-matter-pm-basics.

[6] Yang, Xingchuan, et al. "Comparison of Ground-Based PM2.5 and PM10 Concentrations in China, India, and the U.S." *MDPI*, J. Environ. Res. Public Health, 2 July 2018, www.mdpi.com/1660-4601/15/7/1382/htm.

[7] Mohapatra, Kedar, and S.K. Biswal. "Effect of Particulate Matter (PM) on Plants, Climate, Ecosystems and Human Health." *International Journal of Advanced Technology in Engineering and Science*, 4 Apr. 2014, pdfs.semanticscholar.org/24d4/dfdd07cfbd70bcdb9ef92ec5cc52650cb28a.pdf.

[8] "Basic Information about NO2." *EPA*, Environmental Protection Agency, 8 Sept. 2016, www.epa.gov/no2-pollution/basic-information-about-no2#What%20is%20NO2.

[9] "Chemistry in the Sunlight." *NASA Earth Observatory*, NASA, 28 Jan. 2003, earthobservatory.nasa.gov/features/ChemistrySunlight/chemistry_sunlight3.php.

[10] "How Air Pollution Is Destroying Our Health." *World Health Organization*, World Health Organization, www.who.int/airpollution/news-and-events/how-air-pollution-is-destroying-our-health.

[11] Apte, Joshua S., et al. "Ambient PM2.5 Reduces Global and Regional Life Expectancy." *Environmental Science & Technology Letters*, pubs.acs.org/doi/10.1021/acs.estlett.8b00360.

[12] Simon, Stacy. "World Health Organization: Outdoor Air Pollution Causes Cancer." *American Cancer Society*, 17 Oct. 2013, www.cancer.org/latest-news/world-health-organization-outdoor-air-pollution-causes-cancer.html.

[13] "Nanoparticles and the Lung." *Wissensplattform Nanomaterialien*, www.nanopartikel.info/en/nanoinfo/body-barriers/2388-nanoparticles-and-the-lung.

[14] Morgenstern, Verena, et al. "Atopic Diseases, Allergic Sensitization, and Exposure to Traffic-Related Air Pollution in Children." *American Journal of Respiratory and Critical Care Medicine*, 15 June 2008, www.atsjournals.org/doi/full/10.1164/rccm.200701-036OC.

[15] Weinmayr, Gudrun, et al. "Long-Term Exposure to Fine Particulate Matter and Incidence of Type 2 Diabetes Mellitus in a Cohort Study: Effects of Total and Traffic-Specific Air Pollution." *Environmental Health*, BioMed Central, 19 June 2015, www.ncbi.nlm.nih.gov/pubmed/26087770.

[16] Su, Wei, et al. “The Short-Term Effects of Air Pollutants on Influenza-like Illness in Jinan, China.” *BMC Public Health*, 21 Oct. 2019, www.ncbi.nlm.nih.gov/pubmed/31638933.

[17] Zhang, Zilong, et al. “Particulate Matter Air Pollution and Blood Glucose in Children and Adolescents: A Cross-Sectional Study in China.” *Science of The Total Environment*, 12 July 2019, www.sciencedirect.com/science/article/pii/S0048969719332930.

[18] Zhang, Liqiang, et al. “Air Pollution-Induced Missed Abortion Risk for Pregnancies.” *Nature Sustainability*, Sept. 2019, www.researchgate.net/publication/336020454_Air_pollution-induced_missed_abortion_risk_for_pregnancies.

[19] Yitshak-Sade, Maayan, et al. “Estimating the Causal Effect of Annual PM2.5 Exposure on Mortality Rates in the Northeastern and Mid-Atlantic States.” *Environmental Epidemiology (Philadelphia, Pa.)*, 13 Aug. 2019, www.ncbi.nlm.nih.gov/pubmed/31538134.

[20] Jerrett, Michael, et al. “Spatial Analysis of Air Pollution and Mortality in California.” *American Journal of Respiratory and Critical Care Medicine*, 1 Sept. 2013, www.ncbi.nlm.nih.gov/pubmed/23805824.

[21] Cross, Carroll E., et al. “Ozone Interactions with Lung Tissue .” *The American Journal of Medicine*, June 1976, www.amjmed.com/article/0002-9343(76)90562-3/pdf.

[22] Kraft, Martin, et al. “The German View: Effects of Nitrogen Dioxide on Human Health--Derivation of Health-Related Short-Term and Long-Term Values.” *International Journal of Hygiene and Environmental Health*, U.S. National Library of Medicine, 2005, www.ncbi.nlm.nih.gov/pubmed/16078645.

[23] Thiering, Elisabeth, et al. “Associations of Residential Long-Term Air Pollution Exposures and Satellite-Derived Greenness with Insulin Resistance in German Adolescents.” *Environmental Health Perspectives*, National Institute of Environmental Health Sciences, Aug. 2016, www.ncbi.nlm.nih.gov/pubmed/26863688?dopt=Abstract.

[24] Schneider, Alexandra, et al. “Quantifizierung Von Umweltbedingten Krankheitslasten Aufgrund Der Stickstoffdioxid-Exposition in Deutschland.” *Umweltbundesamt*, 5 Mar. 2018, www.umweltbundesamt.de/publikationen/quantifizierung-von-umweltbedingten.

[25] Von Mutius, E, et al. “Air Pollution and Upper Respiratory Symptoms in Children from East Germany.” *European Respiratory Society*, European Respiratory Society, 1 May 1995, erj.ersjournals.com/content/8/5/723.short.

[26] Chen, Bingheng, et al. “Exposures and Health Outcomes from Outdoor Air Pollutants in China.” *Toxicology*, , 9 Apr. 2004, www.sciencedirect.com/science/article/pii/S0300483X04001040.

[27] Millman, Alexander, et al. “Air Pollution Threatens the Health of Children in China.” *American Academy of Pediatrics*, American Academy of Pediatrics, 1 Sept. 2008, pediatrics.aappublications.org/content/122/3/620.

[28] Zhao, Wenchang, et al. “Urban Ambient Air Quality Investigation and Health Risk Assessment during Haze and Non–Haze Periods in Shanghai, China.” *Atmospheric Pollution Research*, 6 Jan. 2016, www.sciencedirect.com/science/article/pii/S130910421530372X.

[29] Maji, Kamal Jyoti, et al. “Ozone Pollution in Chinese Cities: Assessment of Seasonal Variation, Health Effects and Economic Burden.” *Environmental Pollution*, U.S. National Library of Medicine, Apr. 2019, www.ncbi.nlm.nih.gov/pubmed/30721870.

[30] Westervelt, D M, et al.“Mid-21st Century Ozone Air Quality and Health Burden in China under Emissions Scenarios and Climate Change.” *IOPScience*, 16 July 2019, iopscience.iop.org/article/10.1088/1748-9326/ab260b#erlab260bs4.

[31] Adams, Dara R, et al. “Nitrogen Dioxide Pollution Exposure Is Associated with Olfactory Dysfunction in Older U.S. Adults.” *International Forum of Allergy & Rhinology*, U.S. National Library of Medicine, Dec. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC5554588/.

[32] Fann, Neal, et al. “Estimating the National Public Health Burden Associated with Exposure to Ambient PM2.5 and Ozone.” *Risk Analysis*, U.S. National Library of Medicine, Jan. 2012, www.ncbi.nlm.nih.gov/pubmed/21627672.

[33] Bell, Michelle L, et al. “Ozone and Short-Term Mortality in 95 US Urban Communities, 1987-2000.” *JAMA*, U.S. National Library of Medicine, 17 Nov. 2004, www.ncbi.nlm.nih.gov/pubmed/15547165.

[34] “Meilenstein Für Modernisierung Deutschlands.” *Die Bundesregierung*, 14 Nov 2016, www.bundesregierung.de/Content/DE/Artikel/2016/11/2016-11-14-klimaschutzplan-2050.html.

[35] Mathews, John, et al. “The Spectacular Success of the German Energiewende and What to Do Next.” *Energy Post*, 10 Oct. 2017, energypost.eu/the-spectacular-success-of-the-german-energiewende-and-what-needs-to-be-done-next/.

[36] “Indicator: Greenhouse Gas Emissions.” *Umweltbundesamt*, 25 Apr. 2019, www.umweltbundesamt.de/en/indicator-greenhouse-gas-emissions#textpart-1.

[37] Meyer, Robinson. “Most Americans Support Staying in the Paris Agreement.” *The Atlantic*, Atlantic Media Company, 31 May 2017, www.theatlantic.com/science/archive/2017/05/most-americans-support-staying-in-the-paris-agreement/528663/.

[38] “Strengthening the EPA's Clean Power Plan.” *Union of Concerned Scientists*, 10 Oct. 2014, www.ucsusa.org/resources/strengthening-epas-clean-power-plan.

[39] Roberts, David. “How Trump Policy Will Affect US Carbon Emissions, in One Graph.” *Vox*, Vox, 31 May 2017, www.vox.com/energy-and-environment/2017/5/25/15689200/trump-policy-carbon-emissions-graph.

[40] Sullivan, Timothy J., et al. “Air Pollution Success Stories in the United States: The Value of Long-Term Observations.” *Environmental Science & Policy*, 15 Mar. 2018, www.sciencedirect.com/science/article/pii/S1462901117312352.

[41] Chen, Bingheng, et al. “Air Pollution and Health Studies in China-Policy Implications.” *Taylor & Francis*, 31 Oct. 2011, www.tandfonline.com/doi/full/10.1080/10473289.2011.604288.

[42] Li, Ying, and Ke Chen. “A Review of Air Pollution Control Policy Development and Effectiveness in China.” *IntechOpen*, IntechOpen, 1 Aug. 2018, www.intechopen.com/books/energy-management-for-sustainable-development/a-review-of-air-pollution-control-policy-development-and-effectiveness-in-china.

[43] “China Produces and Consumes Almost as Much Coal as the Rest of the World Combined - Today in Energy.” *US Energy Information Administration*, 14 May 2014, www.eia.gov/todayinenergy/detail.php?id=16271.

[44] Huang, Jing, et al. “Health Impact of China's Air Pollution Prevention and Control Action Plan: an Analysis of National Air Quality Monitoring and Mortality Data.” *The Lancet Planetary Health*, July 2018, www.thelancet.com/journals/lanplh/article/PIIS2542-5196(18)30141-4/fulltext.

[45] “China's President Xi Says Will Continue Years-Long War on Smog.” *Reuters*, Thomson Reuters, 19 Oct. 2017, www.reuters.com/article/us-china-congress-pollution/chinas-president-xi-says-will-continue-years-long-war-on-smog-idUSKBN1CN0CI.

[46] “China's Hebei Province to Cut Coal Use by 5 Million Tonnes This Year: Xinhua.” *Reuters*, Thomson Reuters, 27 Jan. 2018, www.reuters.com/article/us-china-coal-hebei/chinas-hebei-province-to-cut-coal-use-by-5-million-tonnes-this-year-xinhua-idUSKBN1FG0EO.

[47] Rathi, Akshat. “Sweden Legally Commits to Reaching Net-Zero Emissions by 2045.” *Quartz*, Quartz, 16 June 2017, qz.com/1007833/swedens-climate-act-legally-commits-the-country-to-reach-net-zero-emissions-by-2045/.

[48] Wreglesworth, Rob. “17 Innovative Ways to Reduce Air Pollution in Our Cities.” *Disruptive Environmentalist*, 13 June 2019, disruptiveenvironmentalist.com/17-innovative-ways-to-reduce-air-pollution-in-our-cities/.

Authorship Attribution of 50 Victorian Era Novelists with Convolutional Neural Networks

Kei Yamashita,
American School in Japan, Tokyo, Japan
Email: 21yamashitak@asij.ac.jp

Abstract - **Contrary to common methods of authorship attribution, by implementing a multi convolutional layer structure, we are able to reduce training time drastically while maintaining relatively high accuracy. Despite all this, the training of the model was able to be completed whilst implementing a relatively low power computer with a single GPU in just a few hours. The network is able to achieve accuracies of __ in the Victorian Era Novelists Authorship Attribution Dataset.**

Key Words - Convolutional Neural Network, Activation Functions

Introduction

I. Background

In a world of increasing accessibility to information and anonymity, writing in an academic setting has seen a huge impact with plagiarization, and ghostwriters. Not only does this allow people to submit disingenuous works of literature, but lowers the diversity in quality and style of writing. In the educational space, services such as turnitin.com allow institutions to cross-reference academic works against their expansive database of other academic works as well as the internet. However, these services are limited in their ability to merely detect blatant plagiarization. A limitation easily overcome by swapping out vocabulary and mixing the structure of the text. Furthermore, these services are unable to distinguish when works submitted are written by others.

To overcome these issues, these services must be able to discriminate works not solely on whether works have been duplicated, but on patterns of writing styles and linguistic attributes.

In this work, the experiment seeks to measure the ability of neural networks to identify authors with machine learning based on their writing style. CNN's (Convolutional Neural Networks) has been established as a major approach to machine learning problems in general. However, in the realm of natural language processing, the utilization of convolutional neural networks have been largely overshadowed by the stability and accuracy of recurrent neural networks, [5] long-short term memory, [5] and gated recurrent neural networks. This is why the benefits of training on CNNs have been largely overlooked despite its high efficiency and low capital requirements. This experiment games with the use case of consumer-oriented applications of the network where speed and ease of use are key, areas in which CNN's shine.

II. Related Works

In the world of NLP (natural language processing) with machine learning have utilized many methods to solve a range of problems. A popular approach to NLP is the use of attention mechanisms. [1] Attention makes the network focus on a subset of specific details of an input which allows the network to better identify patterns for training. However, for the sake of the relatively small dataset, such mechanisms proved to be redundant in creating results. As for other approaches, the use of recurrent neural networks as stated previously has found its way into many applications of NLP. [2] The use of recurrent neural networks, which generate and retain weights of inputs to then build upon later in training. Such approaches have found itself to be especially useful in the realm of machine translation, where short texts are commonplace.

However, the specifics of this experiment, notably the fact that our training sequences are each 1,000 words long disables the RNN to fully appropriate its strengths, as the long sequence makes it impossible to utilize weight learned early in the sequence, providing no benefit.

III. Advantages of This Work

The usage of CNNs is able to tackle the limitations presented above thanks to its ability to process many texts in parallel, unlike the linear progression of an RNN.

Methodology

I. Equipment

This experiment primarily involves the utilization of machine learning algorithms built on neural networks. A neural network has two phases. First is its learning stage, in which it takes an input—in this case, text data—and runs it through the network to learn values that serve as characteristics of the piece of writing. In the second phase, the network again takes an input, but this time runs the text through the network's learned values to attribute the data to the proper output, enabling the network to identify the author who wrote that piece of text.

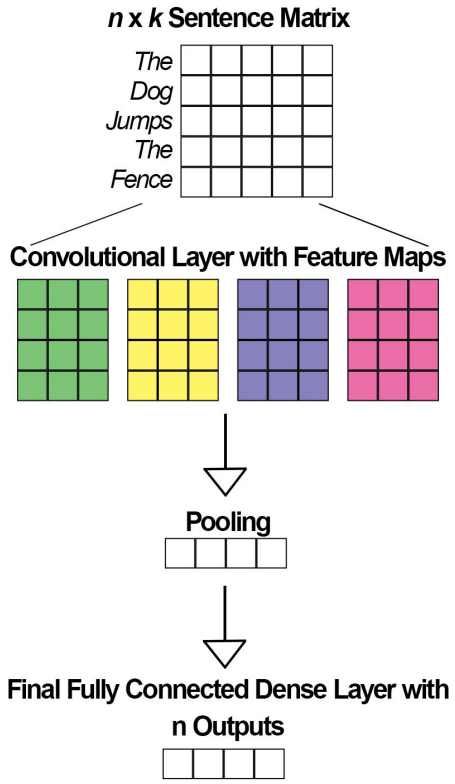


FIGURE I
BASIC CNN STRUCTURE

A convolutional neural network simply put is constructed of several layers of convolutions with non-linear activation functions, [4] such as ReLU or SoftMax functions applied to the outputs. Typically, the convolution is applied over an image, represented in a matrix. However, in the case of our scenario here with text, we represent sentences as a matrix with values assigned to each word in the text through a process called tokenization. With this, the network is able to run convolutions over text to learn vectors and cluster words to learn writing characteristics for each author.

Of course, a network is no use without data to run through it. In this experiment, we will be using the Victorian Era Novelists Authorship Attribution Dataset, [6] created by Abdulmecit Gungor, which consists of 50 authors from the victorian era with texts from over 1000 works. Although this is a relatively small dataset, it has plenty of data to allow the network to distinguish features and discriminate authorship

As for the hardware, the network is trained using a desktop computer with a single Nvidia GTX 1070 graphics card and an Intel Core i7 CPU, which provides the model with plenty of processing power for the network to utilize.

Finally, to prepare for experimentation, the data had to be preprocessed. The text was tokenized [10] with each word having assigned an integer. However, the number of words to have assigned values were capped at 20,000 to maintain the speed and efficiency of the model. Furthermore, in order since our neural networks only accept consistent sequence lengths, each text sequence was "padded," or truncated at 1000 words each.

Once we train the model, the network needs additional data to reference and validate upon to measure how accurate the model is at classifying attributes. Therefore, we split a third of the data to be reserved for testing to be used later. The remaining 2/3rds is now the training data. [8]

II. Experimentation

To arrive to the conclusion of using CNNs, we experimented with a multitude of models and networks in order to derive the best method.

One of the first attempts, the double LSTM [5] method employed the use of two LSTM layers (a type of RNN network) in the model. The structure is seen here.

TABLE I
STRUCTURE OF DOUBLE LSTM NETWORK

Layer Type	Shape
Embedding	(8)
LSTM 1	(150)
LSTM 2	(50)
Dense	(50)
Output	(50)

The embedding layer [9] processes to input to represent each letter with an integer. Then, the data flows through two LSTM layers to learn weights. Finally, the network is parsed into a fully connected dense layer as the output to classify the author. This was a very simple first step and provided accuracies of around 13% consistently. Although that is much better than the 2% odds of random chance, 13% is too low to provide any real-world benefits.

In the next attempt, we implement a Bi-Directional [12] LSTM approach. The bi-directional LSTM consists of two hidden layers which both move linearly along the text to learn weights, however, move in opposite directions, to learn from both past and future states.

TABLE II
STRUCTURE OF BI-DIRECTIONAL LSTM NETWORK

Layer Type	Shape
Embedding	(1000, 8)
Bi-Directional	(1000, 40)
Dense (Time Distributed)	(1000, 1)
Flatten	(1000)
Dense	(50)
Activation	(50)

As seen above, the network utilizes a similar structure to the previous method, however, it replaces the two LSTM layers with a single bi-directional layer as well as introducing a Time Distributed Dense layer as well as an Activation layer at the output. The Time Distributed layer is essentially a dense layer with identical weights and biases in each node. The Activation layer simply serves to be the "gate" that activates when a certain threshold is met, allowing the network to classify an output.

This approach had a significant improvement on the previous attempt, with accuracies up to 31%

For the third attempt, we took the same network from the Bi-Directional LSTM network and added a single convolutional layer after the LSTM layer. We will ignore the specifics of this attempt as the outcomes were very poor, with accuracies regressing back to 14%. However, further experimentation with reducing neuron counts showed significant improvements to the network, giving insight to the nature of the network being prone to overfitting [7] the model. A situation in which a network discovers patterns that only exist in the specific data it is trained on, failing to apply its learned values onto testing data later on despite the input having similar qualities.

This shed light on a weakness of the experiment, in which the small size of the dataset made the model prone to overfitting.

With all this in mind, our experimentation lead us to this model, the multi CNN model, for our final presentation

TABLE III
STRUCTURE OF NEURAL NETWORK LAYERS

Layer Type	Shape
Embedding	(1000, 8)
Convolution 1	(998, 16)
Batch Normalization 1	(998, 16)
Activation 1	(998, 16)
Dropout 1	(998, 16)
Convolution 2	(498, 16)
Batch Normalization 2	(498, 16)
Activation 2	(498, 16)
Dropout 2	(498, 16)
Convolution 3	(166, 16)
Batch Normalization 3	(166, 16)
Activation 3	(166, 16)
Dropout 3	(166, 16)
Convolution 4	(55, 16)
Batch Normalization 4	(55, 16)
Activation 4	(55, 16)
Dropout 4	(55, 16)
Convolution 5	(18, 16)
Batch Normalization 5	(18, 16)
Activation 5	(18, 16)
Dropout 5	(18, 16)
Flatten	(288)
Dense	(50)
Batch Normalization 6	(50)
Output	(50)

As seen in Table III, the basic structure of the network model consists of five convolutional networks which we will call "blocks." Each block has a Batch Normalization, Activation, and Dropout layer. Specifically, the convolutional layers in each block consist of 16 neurons. The decision to only use 16 neurons for each convolution is derived from the fact that as the dataset makes any model prone to overfitting, the model is best trained with a relatively low node count over multiple layers. Furthermore, the convolutions are specified with strides ranging from one to three. Although the first block is set with a stride of one, later blocks have larger strides of two to three to reduce the total number of parameters allowing for faster training. The activation layer associated with each block is a ReLU (rectified linear unit) function, and the dropout is set at 10% excluding the final block, which is set at 50%.

After the five blocks, the outputs are passed through a flatten layer to be routed into a fully connected dense layer with 50 neurons, representative of the 50 authors.

As for other hyperparameters, the learning rate of the model was set at 0.001. This value was drawn after numerous trials and errors and proves to consistently have the highest accuracy for the model.

The actual training of the model is carried out with a batch size of 128 and 50 epochs. It was found that the model no longer found author characteristics and capped its accuracy at around 50 epochs.

RESULTS

With training times of 4 seconds per epoch, the full training of the model was completed in just under 3 minutes. The model showed accuracies of up to 75%—a significant improvement over any other method.

COMPARATIVE ANALYSIS

As stated numerous times before, the limitations of the experiment created by the size of the dataset created numerous obstacles. For the reasoning behind why the last 25% of accuracy is so hard to achieve, as we found lies in the unbalanced distribution of data to each author.

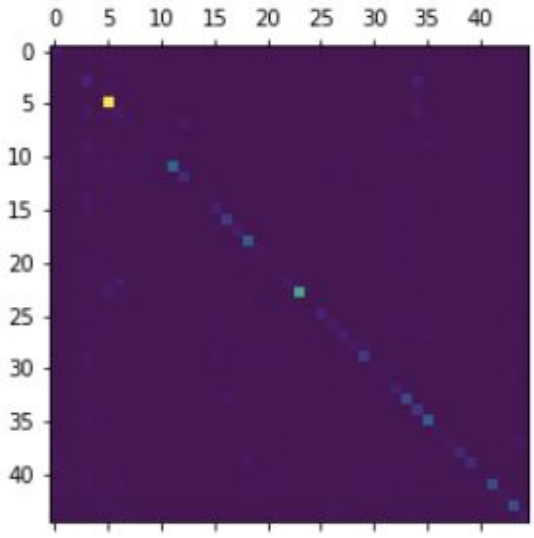


FIGURE 2
DISTRIBUTION OF LEARNED CHARACTERISTICS OF AUTHORS

As seen in Figure 2, there is an imbalance in the distribution of attributed learned within the model. The diagonal line represents the clarity of attributes learned, which brightness of the square indicating how strong characteristics are learned. As you can see, not all authors receive the same amount of attribution. A problem we later found to be rooted in the fundamental imbalance of the dataset, where some authors were favored with more data than others.

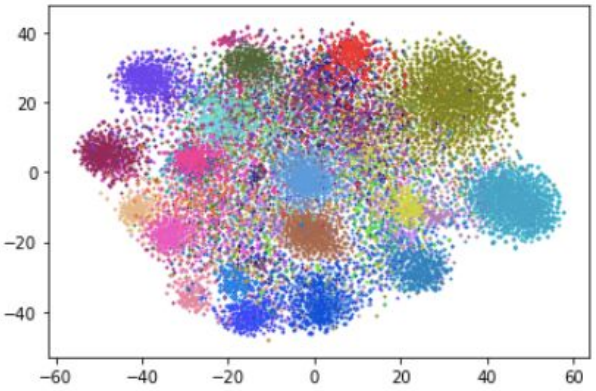


FIGURE 3
TSNE [11] PLOT OF EXTRACTED HIGH DIMENSIONAL FEATURES OF AUTHORS

This is elaborated in Figure 3, where although many features are extracted as seen with clumps of points in the TSNE plot above, the majority of points are dispersed amongst each other with no distinct features to differentiate them from another, posing a difficulty for the network to classify authors.

Hopefully, in future experimentation, a more balanced dataset could be presented to eliminate the issue, or a new methodology can be implemented to overcome disparities in data amounts.

CONCLUSION

In this model, we presented the use of convolutional neural networks in a natural language processing setting that most commonly employs recurrent neural networks. For attribution tasks, the model is able to achieve a state of the art performance while maintaining significantly low training times for the model. The ability to quickly identify patterns and characteristics of an author opens up multitudes of possibilities for many applications in the writing space. In the future, we hope to expand the scope of attribution not just to authorship, but to the quality of writing.

ACKNOWLEDGMENTS

I wish to thank James Borg for his help in mentoring the process of the experiment as well as proof-reading this paper.

REFERENCES

You will list your references like shown below. If there are more than three authors, provide up to three authors' names; replace the others by "*et al*."

[1] Vaswani, A; Shazeer, N; Parmar, N; *et al* "Attention is All You Need," from https://arxiv.org/pdf/1706.03762.pdf

[2] Parmar, M; Devi, V "Neural Machine Translation with Recurrent Highway Networks" from https://arxiv.org/pdf/1905.01996.pdf

[3] Abadi, M; Agarwal, A; Barham, P; et al "TensorFlow: Large-scale machine learning on heterogeneous systems," from https://www.tensorflow.org

[4] Sharma, S "Activation Functions in Neural Networks" from https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

[5] Sherstinsky, A "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network" from https://arxiv.org/pdf/1808.03314.pdf

[6] Gungor, A.” Fifty Victorian Era Novelists Authorship Attribution Data” IUPUI University Library, from http://dx.doi.org/10.7912/D2N65J

[7] Srivastava, A; Hinton, G; Krizhevsky, A; *et al* “Dropout: A Simple Way to Prevent Neural Networks from Overfitting” from http://jmlr.org/papers/volume15/srivastava14a.old/srivastava14a.pdf

[8] Reitermanová, Z “Data Splitting” from https://www.mff.cuni.cz/veda/konference/wds/proc/pdf10/WDS10_105_i1_Reitermanova.pdf

[9] Khrulkov, V; Hrinchuk, O; Mirvakhabova, L; *et al* “Tensorized Embedding Layers for Efficient Model Compression” from https://arxiv.org/pdf/1901.10787.pdf

[10] Hiraoka, T; Shindo, H, Matsumoto, Y; “Stochastic Tokenization with a Language Model for Neural Text Classification” from https://www.aclweb.org/anthology/P19-1158.pdf

[11] Maaten, L; Hinton, G; “Visualizing Data Using t-SNE” from http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf

[12] Schuster, M; Paliwal, K; “Bi-Directional Recurrent Neural Networks” from https://pdfs.semanticscholar.org/4b80/89bc9b49f84de43acc2eb8900035f7d492b2.pdf

IYRC 2020

3rd International Young Researchers' Conference

Important Dates

Submission Deadline: **September 15th**
Notification: **September 30th**
Final Paper Submission: **October 15th**
Registration: **October 15th**
Conference: **November 27th & 28th**

Submission Guidelines

Submitted papers must use the two-column MS-WORD templates provided on our website and not exceed a maximum of six, A4-sized pages. All papers must be submitted in the English language.

The first author of a paper must be a current high school student. The research described must be original and done in whole or substantially by the student(s). Submitted papers will undergo a blind review process. Authors of accepted papers will be invited to give oral or poster presentations. All papers will be edited and published.

Topics

Science
Math
Education
Humanities
Literature
Sociology
Commerce

Registration

Authors of accepted papers will be required to register and attend the conference to present their papers. Non-authors and non-presenters may register under general registration. Author registration covers conference entry, lunch, dinner and a single copy of the proceeding. General registration also covers conference entry, lunch, dinner and a single copy of the proceeding.

Awards

Awards for best papers and presentations will be given.

Stay Connected

URL: **https://the-iyrc.org/**
Email: **info@the-iyrc.org**

Nov 27-28, 2020

Call for Papers

The organizing committee is pleased to announce a Call for Papers for the 3rd International Young Researchers' Conference.

The goal of this conference is to provide an avenue for budding scientists and motivated high school students to share and present their school research results with their peers and the wider community. It is intended to help foster the spirit of research in high schools. The two-day conference will contain additional activities such as keynote speeches from eminent teachers, scholars and industry professionals.

Take a picture to visit the IYRC website