



ISF 弘立

BAUHINIA

紫荊花



The Student Research Journal of The ISF Academy
弘立書院學生研究期刊

Volume VII
Issue 2, 2022

Editors: *Ms. C. Brillaux, Mr. J. Faherty, Dr. L. Gao, Dr. S. D. J. Griffin,*
Ms. D. Ibarra, Mr. K. Kampen, Dr. R. Oser, Dr. M. Pritchard, Dr. F. Saunders,
Ms. D. Wang, Dr. H. Y. Wu, Mr. F. Wynne, Ms. D. Yeung, Dr. P. Yuen, Dr. Y. L. Zhang

ISSN 2409-4064

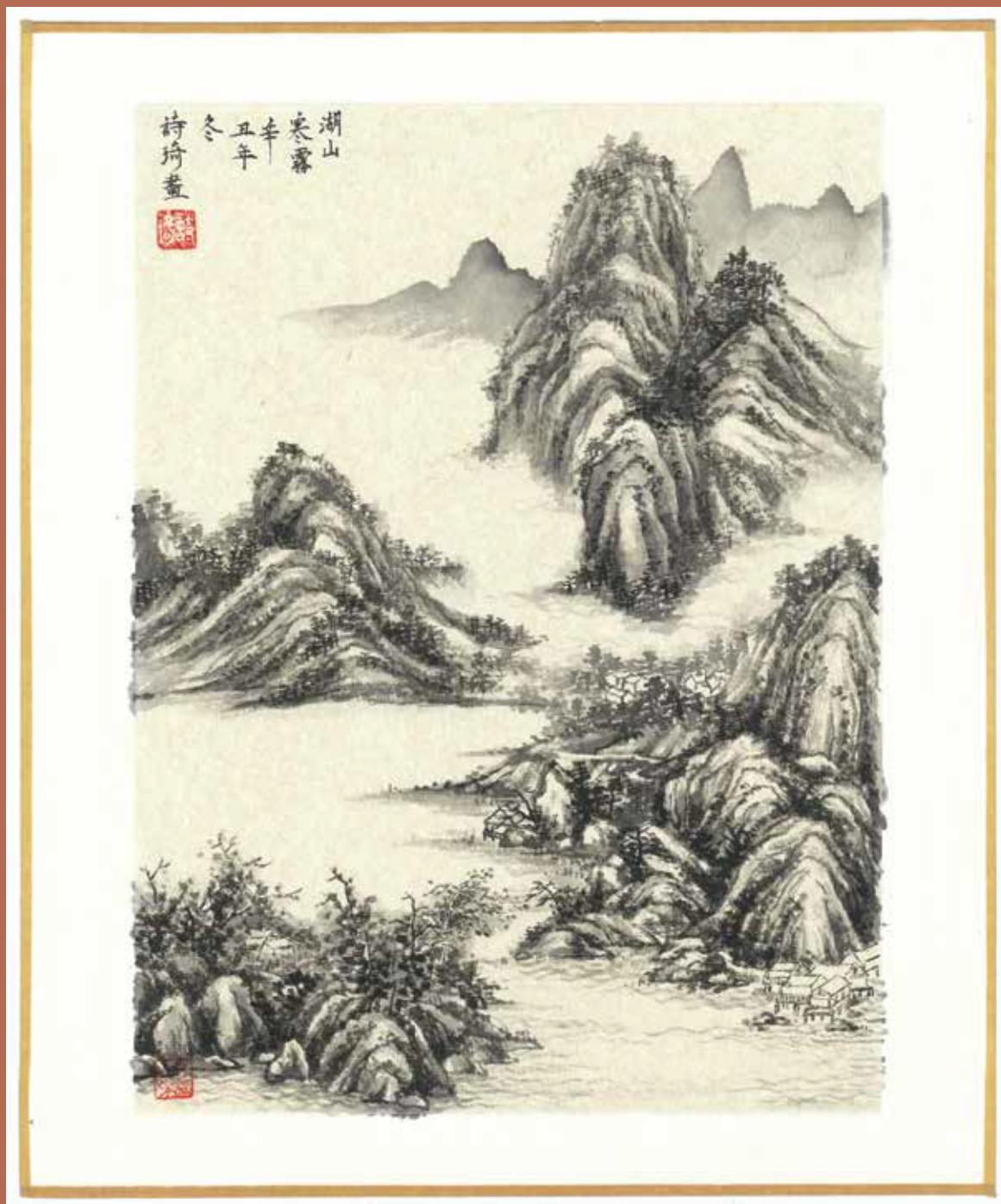


Table of Contents 目錄

The antimicrobial effects of ethanol	1
<i>Chrystal Li 李瑩瑩</i>	
“Statistics conceal as much as they reveal.”	17
<i>Denton Philtjens 費丹成</i>	
Fourier Series modelling with applications to electrocardiogram readings	21
<i>Dionne Daiyin Yeung 楊岱殷</i>	
Design Technology: EpiPen	33
<i>Stanley Ip 葉宇軒</i>	
Hybrid Sentiment Analysis system to extract language bias in news media	65
<i>Sally Sijie Song 宋思婕</i>	
Within areas of knowledge, how can we differentiate between change and progress?	74
<i>Candace Yan Yue Chung 鍾欣瑜</i>	
Smallpox and its associated political implications in Qing dynasty China	78
<i>William Tristan Lee III 李欣</i>	
A data-driven study on the effects of changes in local sea surface temperature on the health of Coral Reefs in Hoi Ha Wan, Hong Kong	84
<i>Julie Tam 譚幸臨</i>	
A study on the relationship between the colour and mass of selected main sequence stars	99
<i>Candace Yan Yue Chung 鍾欣瑜</i>	
Proving Fermat’s theorem on the sum of two squares by elaborating on Zagier’s “one-sentence” proof	110
<i>Lok Tong Coco Yeung 楊樂同</i>	
How effective is online learning during the COVID-19 pandemic, according to student’s perceptions?	117
<i>Andrew Minghan Jiang 蔣明翰</i>	
Modelling COVID-19 using the SIHD model	124
<i>Ying Chun Justin Man 文言中</i>	

The title of the painting is “Cold Fog on Lake and Mountain”. The intention is to depict a misty winter mountain scene, creating a poetic world where “it is possible to work, to look, to travel, and to live” (in the words of Guo Xi, a painter of the Northern Song Dynasty). At the bottom of the painting, boats pass by without a trace, leaving only ripples of water of varying depths, inviting one’s imagination; several groups of rustic huts and winding mountain paths loom between rocks and trees, waiting to be explored; in the distance, misty mountain colors and a hint of cold greenery indicate the arrival of spring. This painting was completed on New Year’s Day, 2022.

畫的標題為《湖山寒霧》，黃詩琦意在描繪一個煙霧迷濛的冬日山色圖景，創造一個“可行、可望、可遊、可居”（北宋畫家郭熙提出）的詩意世界。畫面下方船隻駛過不見蹤影，只留下深淺不一的水波紋，引人遐想；幾座古樸茅屋和盤旋山路在山石和樹林之間若隱若現，待人探索；遠處迷濛的山色和點點寒翠昭示著春的來臨。此畫完成於 2022 年元旦。

Editor's Note

It is with great pleasure and pride that we present the seventh edition of Bauhinia. Through an ongoing pandemic and global weariness, our students continue to astound us with their commitment to academic endeavors.

Bauhinia aims to honor this excellent scholarship, emerging either from our distinguished curriculum at the ISF Academy, and/or as a product of our unique extracurricular programs offered by Shuyuan. The mission to be “Independent, Global, and Chinese” is manifested within the topic choices and self-directed learning that went into these papers. The work ethic, as well as the research and communication skills gained through writing these articles reflect the Eight plus One Virtues espoused at the ISF Academy, namely ‘zhi’ (Intelligence Wisdom) and ‘ai’ (Passion for Learning Life).

The works presented continue to grow in quality and quantity and thus, the Bauhinia is once again offered in two parts: the first issue contains articles that are more Humanities-focused, while the second issue is more focused on STEM (Science, Technology, Engineering, and Mathematics). The diversity of articles also illustrates the bilingual and multidisciplinary nature of our school. Topics herein include how Lucretius and Xunzi’s interpretations of the problem of evil reflect their conceptions of divinity, a study on the relationship between the colour and mass of selected Main Sequence Stars, and an analysis of ‘rain’ imagery in Nalan Xingde’s 納蘭性德 lyrics 詞.

It is imperative to highlight that the layout of this seventh edition of Bauhinia is a product of home-grown talent: Daisy Wang and Dionne Yeung, two students who recently graduated and enthusiastically partook in the editing process. We thank the entire Editorial Board for their time and dedication in reviewing this publication, and it is our hope that such intellectual pursuits are what will continue to bind us as students, mentors, editors and designers within a truly collaborative academic community.

You are invited to join this community by exploring the content that sparked such passion in our students. Your engagement and responses are most welcome: sy_team@isf.edu.hk.

Rachel Oser

關於文體的說明

我們非常高興和自豪地推出第七期《紫荊花》。在全球疫情之下，許多人都身心俱疲，而我們的學生繼續以他們對學術研究的投入，給我們帶來驚喜。

《紫荊花》旨在表彰這些優秀的學術成果，這些成果來自於弘立學院的優秀課程，以及「書院」項目所提供的獨特的課外活動。在這些論文的選題和自主研究中，同學們展現了「獨立精神、中華美德、全球視野」的弘立書院使命。通過撰寫這些文章，同學們增進了學術操守以及研究和溝通技巧，體現了弘立書院所倡導的「八德一智」，即「智」（聰明智慧）和「愛」（對學習和生活的熱情）。

同學們的研究論文在質量和數量上都不斷增加，因此，本期《紫荊花》再次分為兩輯：第一輯文章側重於人文科學，而第二輯則更側重於STEM（科學、技術、工程和數學）。文章的多樣性體現了我們學校的雙語和多學科性質。其中的主題包括：盧克萊修和荀子對「惡」問題的解釋如何反映他們對神性的概念，對所選定的主序星的顏色和質量之間關係的研究，對納蘭性德詞中「雨」意象的分析，等等。

必須強調的是，第七期《紫荊花》的版面設計來自於本校人才。王雨涵和楊岱殷，這兩位剛剛畢業的學生，熱情地參與了編輯工作。我們感謝整個編委會為審閱這份出版物所付出的時間和精力，我們希望這樣的智識追求，能夠繼續把我們作為學生、老師、編輯和設計師的不同角色，在一個真正具有合作氛圍的學術社群中聯繫起來。

我們邀請你加入這個社群，探索這些激發出學生們如此的學術熱情的內容。我們熱切歡迎你的參與和回應：sy_team@isf.edu.hk。

歐睿秋

The antimicrobial effects of ethanol

Chrystal Li 李瑩瑩

Research Question

What is the effect of ethanol concentrations (40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%) on the growth of bacteria *Staphylococcus epidermidis*, indicated by the zone of inhibition (mm) on LB agar plates for 24 hours and 48 hours?

Introduction

As of November 3, 2020, there are 47.5 million confirmed SARS-CoV-2 cases worldwide (Worldometer, 2020). The high transmission rates of SARS-CoV-2 (Jefferson *et al.*, 2020) is due to its surface stability, remaining active on surfaces from 24 hours to 9 days (van Doremalen *et al.*, 2020). Within the time interval, coronavirus strains can be transferred by contact with hands on contaminated surfaces. Therefore, hands are considered as the primary route for transmitting coronavirus (Microbiology Society, 2020).

The Centres for Disease Control and Prevention, and the World Health Organization (WHO) promote hand hygiene as the most crucial measure in the prevention of infections, and recommended the use of alcohol-based hand sanitizers (Centers for Disease Control and Prevention, 2019; World Health Organization, 2009). However, their antimicrobial effects only last for two minutes (Black, 2010), which raises questions about whether alcohol only temporarily suppresses microbial growth in the two minutes, or completely inactivates growth in a longer period.

The antimicrobial effects of alcohol-based hand sanitizers are dependent on the alcohol concentration. It is optimum between concentrations of 60% to 85% (McDonnell & Russell, 1999), as added water content slows down the evaporation rate, allowing alcohol to have longer contact with the skin

surface for antimicrobial effects. However, a low alcohol concentration would not sufficiently disrupt the microbes' entire membrane and coagulate all proteins, hence a decrease in effectiveness at concentrations below 50% (McDonnell & Russell, 1999).

This investigation examined the antimicrobial effects of different ethanol (C_2H_5OH) concentrations against the bacteria *S. epidermidis*. Using non-pathogenic *S. epidermidis* to model SARS-CoV-2 was due to the risk of infection when using pathogenic viruses at the school laboratory. Despite the differences between the membrane structure of *S. epidermidis* and SARS-CoV-2, both membranes are composed of phospholipids and proteins, hence are disrupted by ethanol.

1 Background

Ethanol has two antimicrobial mechanisms: phospholipid dissolution and protein denaturation (Ingólfsson & Andersen, 2011; Ophardt, 2003).

Ethanol is freely permeable across microbial membranes (Osman & Ingram, 1985) due to its similar amphipathic properties with phospholipids. The polar phospholipid head is attracted to the polar hydroxyl group of ethanol, while the non-polar phospholipid fatty acid tail is attracted to the non-polar ethyl group of ethanol (Figure 1), leading to phospholipid dissolution (Ingram, 1989). The hydroxyl group of ethanol interferes with the hydrogen bonds in the secondary and tertiary structure of membrane proteins, and forms new hydrogen bonds with the protein side-chain (Ingram, 1989), causing the loss of functionality due to conformational changes in membrane protein structure.

SARS-CoV-2 are enveloped viruses, which contain genetic material packaged in a protein coat,

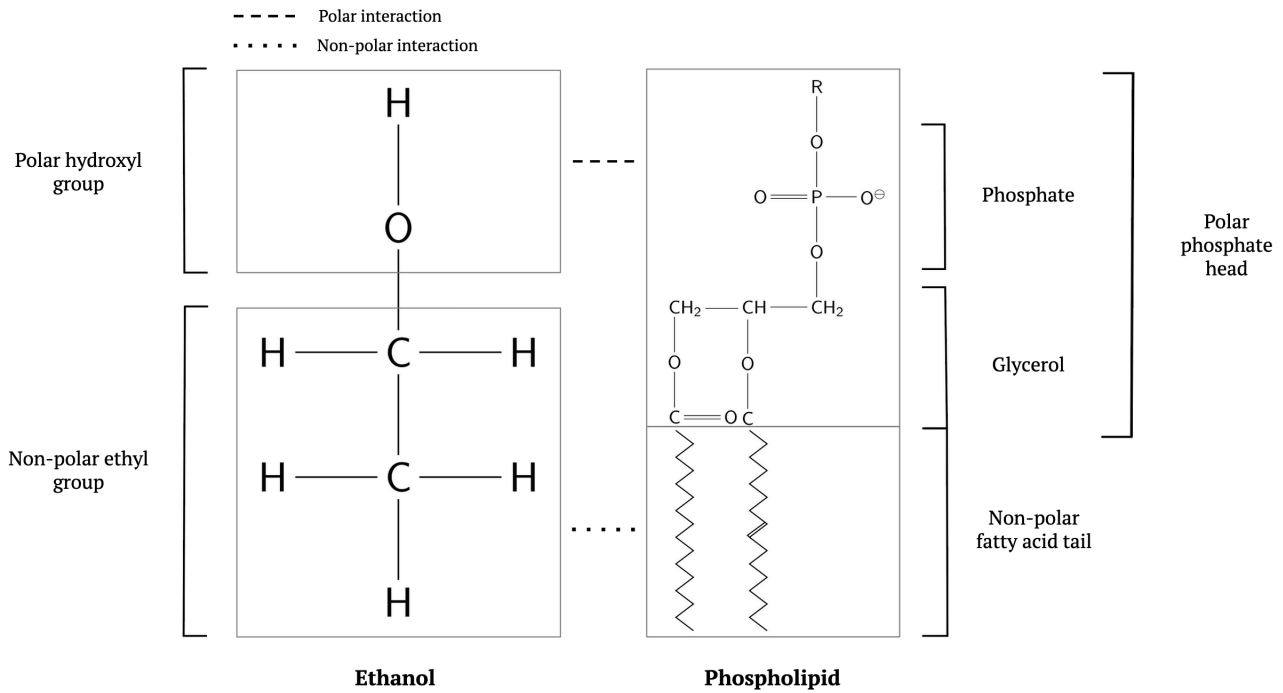


Fig. 1. Attraction between phospholipid and ethanol.

capsid, surrounded by a lipid envelope (Figure 2). These components protect the genetic material, and are crucial for transmitting to another host (Golin, Choi, & Ghahary, 2020). After ethanol's interference with the structure of the protein capsid and lipid envelope, the genetic material is exposed and the viruses cease to survive.

S. epidermidis are gram-positive bacteria, containing a cell wall with a thick layer of peptidoglycan (Figure 3), and an inner plasma membrane that composes of 40% phospholipids and 60% protein (Bruslind, 2020). Ethanol reduces the cell wall's peptidoglycan cross-linking (Cao *et al.*, 2017), thus increasing its permeability for ethanol to reach the membrane. The membrane contains enzymatic proteins that are involved in physiological functions, including solute and electron transport and ATP synthesis (Huffer *et al.*, 2021). Following ethanol's interference with the membrane, the enzymatic proteins are denatured, hence the loss of cellular function.

As ethanol shows the same antimicrobial mechanisms in both microbes, *S. epidermidis* substituted SARS-CoV-2 in this investigation.

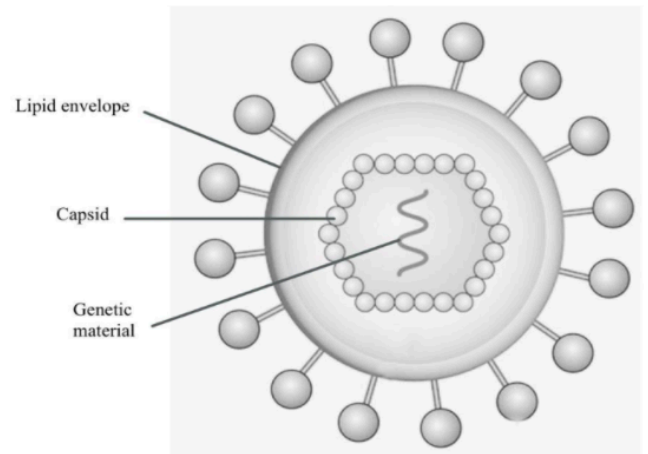


Fig. 2. Membrane structure of enveloped viruses (Wienczek, 2020).

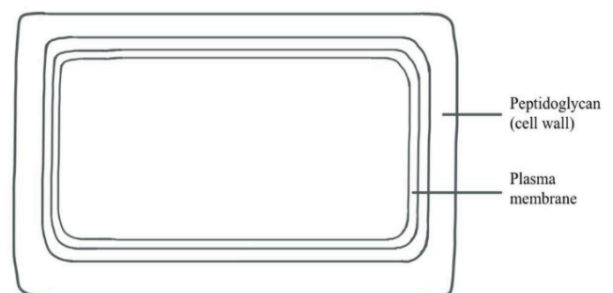


Fig. 3. Membrane structure of gram-positive bacteria (Karki, 2017).

2 Hypothesis

Supported by McDonnell's study, there is a significantly smaller zone of inhibition at ethanol concentrations below 50% (McDonnell & Russell, 1999). At concentrations between 50% to 85%, there is a positive correlation between the two variables, as more ethanol molecules perform bactericidal effects (Ingram, 1989). At 90%, the zone of inhibition decreases due to the evaporation of ethanol (McDonnell & Russell, 1999), reducing its contact time with *S. epidermidis* to perform bactericidal effects.

The predicted trend (Figure 4) is supported by Yamashita's study (Figure 5), where the bacteria's survival rate is high (small zone of inhibition) at low ethanol concentrations; gradually decreases in middle concentrations; increases at high concentrations. However, as Yamashita used a lower incubation temperature (Yamashita & Takeno, 2001), the bacteria's metabolic reactions decrease (Nedwell, 1999), and could be killed with a lower ethanol concentration. Therefore, despite the bacteria having the lowest survival rate at 70% ethanol, 85% ethanol is predicted to have the largest zone of inhibition.

3 Methodology

3.1 Independent Variable

The independent variable was the ethanol concentration (%). Eleven ethanol concentrations: 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, were tested for six trials each to reduce the standard deviation. Though the ideal concentration is between 60% to 85% (McDonnell & Russell, 1999), exposure to high ethanol concentrations is dangerous (French Agency for Food, Environmental and Occupational Health Safety, 2016), hence the range was expanded from the ideal concentration to investigate whether a lower ethanol concentration was capable of inducing similar antimicrobial effects. The ethanol concentrations were manipulated by diluting 100% ethanol with autoclaved water (Table 1).

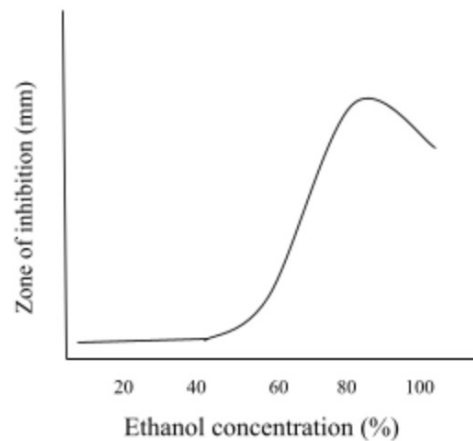


Fig. 4. Sketch graph demonstrating the effect of ethanol concentrations on the zone of inhibition.

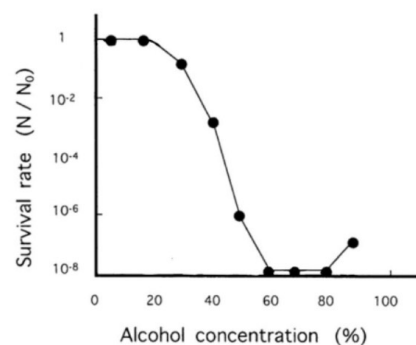


Fig. 5. Effect of ethanol on the survival of *S. aureus* ATCC 12600. The bacterial cells were treated for 10 minutes at 20°C.

3.2 Dependent Variable

Ethanol's bactericidal effects were examined using the disc diffusion method, as microbes must be cultured in closed containers and remain unopened in school laboratories (Microbiology in Schools Advisory Committee, 2010). The dependent variable was the zone of inhibition (mm): the distance between the two opposite points on the zone edge subtracting the diameter of the paper disc. This was measured by a caliper after 24 and 48 hours of incubation. Ideally, time intervals should be decided based on the reproductive rate of *S. epidermidis*. However, quoting Namvar's study, there is a "lack of information about *S. epidermidis* life cycle" (Namvar et al., 2014). Additionally, under COVID-19 pandemic, students were allowed to visit the laboratory once per day, hence the smallest time interval of 24 hours.

Paper discs immersed in different ethanol solutions were placed on agar plates cultured with *S. epidermidis*. The ethanol concentration decreased as it diffused out from the centre of the paper disc on the agar surface. If the ethanol was effective against *S. epidermidis* at a particular concentration, there was no bacterial growth. The region of no bacterial growth was the zone of inhibition, and was in different observed turbidity from other regions of bacterial growth (Figure 6). The larger the zone of inhibition, the greater the bactericidal effect.

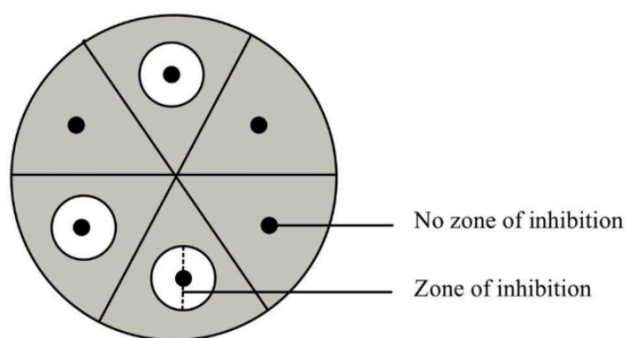


Fig. 6. An agar plate with zone of inhibition marked.

3.3 Control Variables

The following variables were controlled for an accurate basis of comparison:

Type of alcohol: Alcohols have different chemical properties, hence different bactericidal effects. Ethanol was chosen for this investigation, as it is effective against a large spectrum of microorganisms that can linger on the skin (Huynh-Delerme *et al.*, 2012), and is safe from its low dermal absorption (Huynh-Delerme *et al.*, 2012).

Incubation temperature: Temperature affects the enzyme activities in *S. epidermidis*, leading to changes in its metabolism and growth pattern. Despite 37°C being the ideal incubation temperature for *S. epidermidis* (Zaharia *et al.*, 2010), safety precautions of incubating at the human temperature required alternation to 27°C on the incubator.

Starting concentration of *S. epidermidis* on each plate: If the starting concentration of bacteria increases (bacteria are closer to each other), as each ethanol molecule extirpates a fixed number of bacteria, the zone of inhibition decreases and

vice versa. To control, 200 microlitres of *S. epidermidis* was added to each plate, measured by P200 micropipette.

Ethanol molecules added to each paper disc:

It changes the moles of ethanol available to perform bactericidal effects. To control, 5 microlitres of ethanol was added to each paper disc, measured by P20 micropipette.

Agar medium: *S. epidermidis* rely on nutrients in the agar for growth. Different mediums contain different concentrations of nutrients, hence changes in its growth pattern. LB agar plates prepared by school technicians were used for all trials.

3.4 Preliminary Investigation

In the preliminary investigation, paper discs were soaked in Petri dishes with ethanol, and left to dry for 30 minutes. *S. epidermidis* was diluted with LB broth in a 1:1 ratio before adding into the plates. After 48 hours of incubation, all plates were found with no zone of inhibition and indiscernible bacterial growth. After revisiting procedure, 200 microlitres of *S. epidermidis* culture were added into the plates without dilution. Paper discs with 5 microlitres of ethanol were applied directly to prevent the evaporation of ethanol.

3.5 Apparatus List

The materials and equipment needed for this investigation are listed below:

4 mL	<i>S. epidermidis</i> (ATCC®14990™)
10 mL	100% ethanol
10 mL	Autoclaved water
100	6 mm filter paper discs
11	LB agar plates
1	P20 micropipette (±0.1µl)
1	P200 micropipette (±1µl)
1	P1000 micropipette (±5µl)
1	Vortex mixer
1	Plate spreader
1	Tweezer
1	Caliper (±0.5 mm)
11	Parafil
11	1.5 mL microcentrifuge tube

3.6 Procedure

Sterilization

To prevent contamination of *S. epidermidis*, a sterile environment was set up by disinfecting the laboratory bench with 90% ethanol, working within the zone of sterility of a spirit burner, and wearing a laboratory coat and gloves. Tweezers were sterilized by 90% ethanol and flame when transferring the paper discs to the plates.

Dilution

Ethanol concentration (%)	Volume of 100% ethanol (μl)	Volume of autoclaved water (μl)
40	200 ± 1	300 ± 5
45	225 ± 5	275 ± 5
50	250 ± 5	250 ± 5
55	275 ± 5	225 ± 5
60	300 ± 5	200 ± 1
65	325 ± 5	175 ± 1
70	350 ± 5	150 ± 1
75	375 ± 5	125 ± 1
80	400 ± 5	100 ± 1
85	425 ± 5	75 ± 1
90	450 ± 5	50 ± 1

Table 1. Proportion of 100% ethanol and autoclaved water needed for desired ethanol concentrations.

The different uncertainties were due to the different sizes of micropipette used. P200 micropipette ($\pm 1\mu\text{l}$) and P1000 micropipette ($\pm 5\mu\text{l}$) were used when the volume of substance was 200 microlitres or below and above 200 microlitres respectively.

For each concentration, the respective volume of 100% ethanol and autoclaved water were added to a 1.5 mL microcentrifuge tube, and placed on a vortex mixer for a thorough mix.

Plating

200 microlitres of *S. epidermidis* culture were plated on LB agar. The plate was split into six sections with one paper disc placed in the centre of each section (Figure 7). 5 microlitres of ethanol in desired concentration were added directly on the paper disc. The plate was sealed with parafilm and incubated under 27°C. The zone of inhibition for

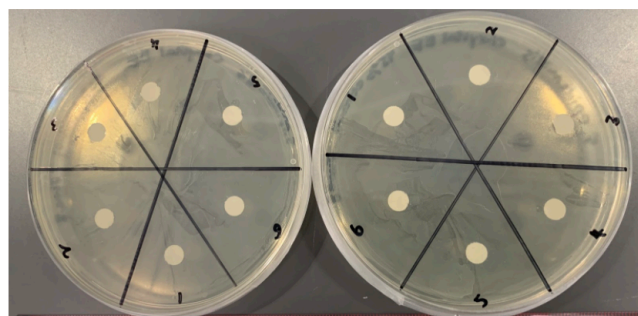


Fig. 7. Setup of agar plates.

each paper disc was measured with a caliper after 24 and 48 hours of incubation.

3.7 Risk Assessment

Bacterial growth: the entry of pathogenic bacteria from the surrounding was a risk for growing harmful microbes in the agar plates. This was prevented by incubating below the human temperature at 27°C, and maintaining a sterile environment (Procedure - Sterilization). The agar plates were sealed with parafilm before incubation, never opened in the process, and were treated as biological waste after the investigation.

Spirit burner & 100% ethanol: avoid storing highly flammable 100% ethanol near the spirit burner to prevent starting a fire. To reduce the risk of burns from accidental contact with hot surfaces, allow any hot apparatus to cool down before handling, and turn off the spirit burner when not in use. Ensure a fire extinguisher is stored in the laboratory in case of a fire.

Plastic waste: To reduce plastic waste, a glass spreader instead of plastic spreaders was used.

4 Data Collection & Processing

4.1 Qualitative Observations

The change in observed turbidity of agar plates indicated bacterial growth. There was consistent turbidity on agar plates with ethanol concentrations between 40% to 55%, hence no detectable zone of inhibition. There were changes in turbidity around the paper discs with ethanol concentration between 60% to 90%, suggesting bacterial inhibition.

4.2 Raw Data

Ethanol concentration (%)	Zone of inhibition (mm) (± 0.5 mm)					
	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6
40	0.0	0.0	0.0	0.0	0.0	0.0
45	0.0	0.0	0.0	0.0	0.0	0.0
50	0.0	0.0	0.0	0.0	0.0	0.0
55	0.0	0.0	0.0	0.0	0.0	0.0
60	2.0	1.0	2.0	0.0	1.0	2.0
65	3.0	1.0	5.0	2.0	3.0	3.0
70	3.0	2.0	2.0	2.0	3.0	2.0
75	3.0	4.0	1.0	4.0	2.0	2.0
80	2.0	5.0	2.0	3.0	3.0	2.0
85	2.0	2.0	5.0	3.0	3.0	4.0
90	4.0	4.0	2.0	3.0	4.0	2.0

Table 2. Mean zone of inhibition vs ethanol concentrations after 24 hours of incubation

Ethanol concentration (%)	Zone of inhibition (mm) (± 0.5 mm)					
	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6
40	0.0	0.0	0.0	0.0	0.0	0.0
45	0.0	0.0	0.0	0.0	0.0	0.0
50	0.0	0.0	0.0	0.0	0.0	0.0
55	0.0	0.0	0.0	0.0	0.0	0.0
60	1.0	0.0	2.0	0.0	1.0	2.0
65	1.0	0.0	4.0	2.0	2.0	1.0
70	2.0	1.0	2.0	2.0	2.0	1.0
75	2.0	4.0	1.0	4.0	2.0	2.0
80	1.0	5.0	2.0	2.0	3.0	2.0
85	0.0	0.0	0.0	0.0	0.0	0.0
90	0.0	0.0	0.0	0.0	0.0	0.0

Table 3. Mean zone of inhibition vs ethanol concentrations after 48 hours of incubation.

4.3 Processed Data

To compare the bactericidal effects of different ethanol concentrations, the mean zone of inhibition was calculated. Mean was chosen over the median, as there were no outliers in the data. Standard deviation was used to measure the spread and variability of data in each ethanol concentration. X-error bars plotted in Figures 8 and 9 show standard deviation for each ethanol concentration.

Calculations for mean zone of inhibition (mm):

$$\frac{\text{zone of inhibition}}{\text{number of trials}}$$

Example (60% ethanol after 24 hours):

$$\frac{2 + 1 + 2 + 0 + 1 + 2}{6} = \frac{8}{6} = 1.3\text{mm}$$

Calculations for standard deviation:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

σ = population standard deviation

N = the size of the population

x_i = each value from the population

μ = the population mean

Ethanol concentration (%)	Mean zone of inhibition (± 0.5 mm)	Standard deviation (2 d.p)
40	0.0	0.00
45	0.0	0.00
50	0.0	0.00
55	0.0	0.00
60	1.3	0.82
65	2.8	1.33
70	2.3	0.52
75	2.7	1.21
80	2.8	1.17
85	3.2	1.17
90	3.2	0.98

Table 4. Zone of inhibition vs ethanol concentrations after 24 hours of incubation.

Ethanol concentration (%)	Mean zone of inhibition (± 0.5 mm)	Standard deviation (2 d.p)
40	0.0	0.00
45	0.0	0.00
50	0.0	0.00
55	0.0	0.00
60	1.0	0.89
65	1.7	1.37
70	1.7	0.52
75	2.5	1.22
80	2.5	1.38
85	2.8	1.17
90	3.2	1.47

Table 5. Zone of inhibition vs ethanol concentrations after 48 hours of incubation.

4.4 Statistical Test

The overlap in error bars suggests insignificant differences in the zone of inhibition between ethanol concentrations, hence two-sample t -tests were conducted under 0.05 significant level. Despite ANOVA giving a lower probability of Type I error than the t -test (Laerd Statistics, 2018), it measures whether there are significant differences between multiple ethanol concentrations. To evaluate the hypothesis, t -test that compares two adjacent concentrations was more appropriate.

Null Hypothesis (H_0): There is no difference between the two ethanol concentrations

Alternative hypothesis (H_1): There is a difference between the two ethanol concentrations

A p -value was given in the two-sample t -test. If the p -value was below 0.05, the null hypothesis was rejected, hence a difference in the zone of inhibition between the two ethanol concentrations. If the p -value was above 0.05, the null hypothesis was accepted, hence no differences between the two ethanol concentrations.

Paired t -tests were conducted under 0.05 confidence level to determine whether there was a significant difference in the zone of inhibition between 24 and 48 hours of incubation, hence exam-

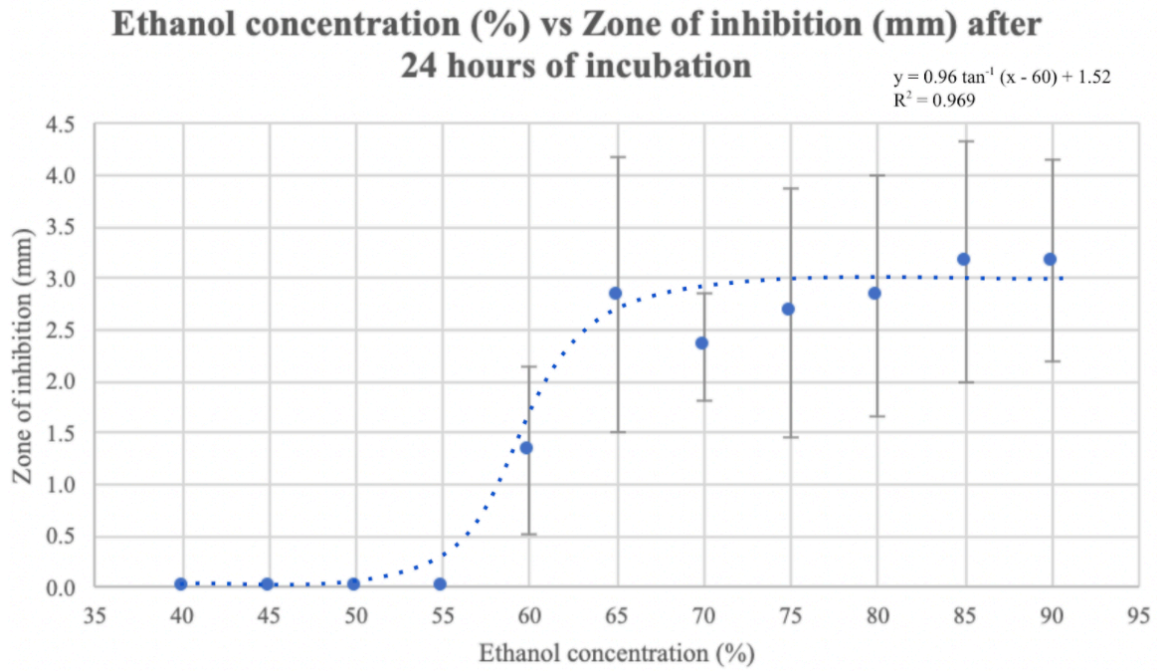


Fig. 8. Effect of ethanol concentrations on the zone of inhibition after 24 hours of incubation.

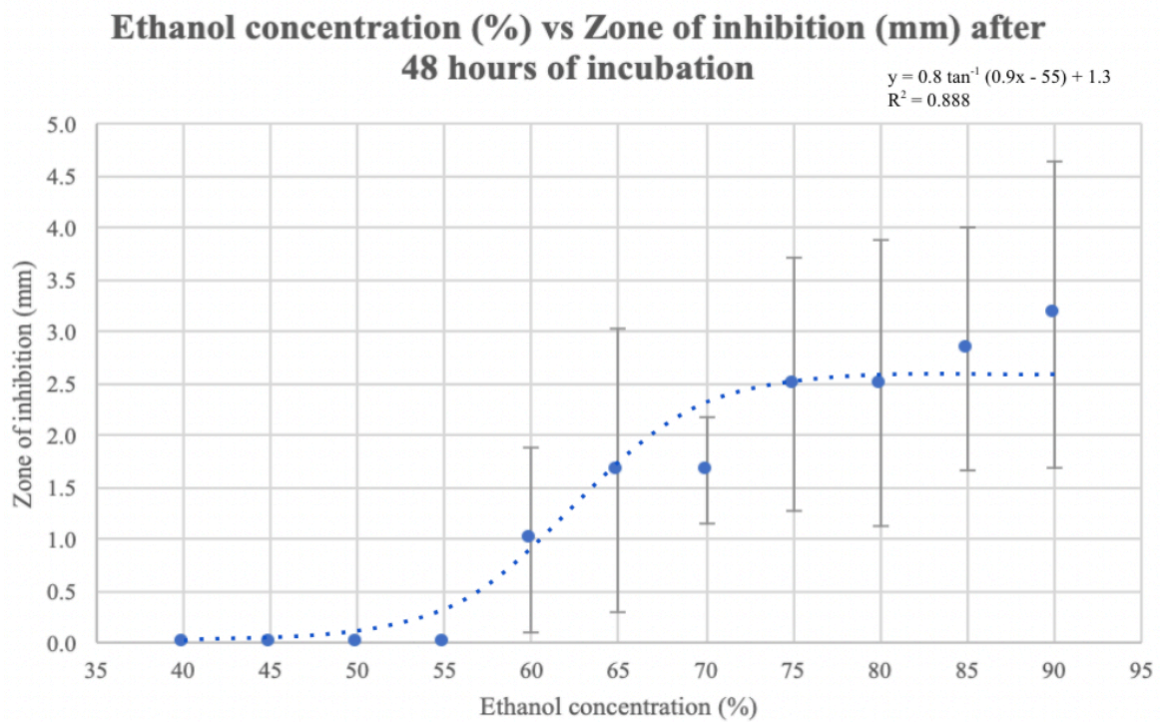


Fig. 9. Effect of ethanol concentrations on the zone of inhibition after 48 hours of incubation.

ining whether ethanol's bactericidal effects reduce over time. Paired *t*-tests were chosen, as the zone of inhibition after both incubation durations were taken from the same samples.

Null hypothesis (H_0): There is no difference between the two incubation durations

Alternative hypothesis (H_1): There is a difference between the two incubation durations

A *p*-value was given in the paired *t*-test. If the *p*-value was below 0.05, the null hypothesis was rejected, hence a difference in the zone of inhibition between the two incubation durations. If the *p*-value was above 0.05, the null hypothesis was accepted, hence no differences between the two incubation durations.

Test description	p-value	T-test Result (Reject H ₀ = R, Accept H ₀ = A)
60% vs 65% (24 hours)	0.046	R
65% vs 70% (24 hours)	0.423	A
70% vs 75% (24 hours)	0.555	A
75% vs 80% (24 hours)	0.813	A
80% vs 85% (24 hours)	0.632	A
85% vs 90% (24 hours)	1.000	A
60% vs 65% (48 hours)	0.343	A
65% vs 70% (48 hours)	1.000	A
70% vs 75% (48 hours)	0.168	A
75% vs 80% (48 hours)	1.000	A
80% vs 85% (48 hours)	0.661	A
85% vs 90% (48 hours)	0.673	A

Table 6. Results of two-sample *t*-tests.

Test description	p-value	T-test Result (Reject H ₀ = R, Accept H ₀ = A)
60% ethanol	0.175	A
65% ethanol	0.013	R
70% ethanol	0.025	R
75% ethanol	0.363	A
80% ethanol	0.175	A
85% ethanol	0.175	A
90% ethanol	1.000	A

Table 7. Results of paired *t*-test.

4.5 Uncertainties

Uncertainties of equipment could lead to anomalies, hence were calculated with the following equation:

Percentage uncertainty:

$$\frac{\text{absolute uncertainty of equipment}}{\text{size of measurement}} * 100\%$$

Percentage uncertainties of ethanol concentrations were determined by the sum of percentage uncertainty for the volume of 100% ethanol and volume of autoclaved water by micropipettes.

Example (40% ethanol):

Percentage uncertainty for volume of 100% ethanol:

$$\frac{1}{200} \times 100\% = 0.50\%$$

Percentage uncertainty for volume of autoclaved water:

$$\frac{5}{300} \times 100\% = 1.67\%$$

Total percentage uncertainty for ethanol concentration:

$$0.50\% + 1.67\% = 2.17\%$$

Same calculations were made for other concentrations, results presented in Table 8.

The percentage uncertainty for volume of *S. epidermidis* and volume of ethanol added to each paper disc remain the same in all trials.

Calculations for percentage uncertainty of volume of *S. epidermidis* by P200 micropipette:

$$\frac{1}{200} \times 100\% = 0.5\%$$

Calculations for percentage uncertainty of volume of ethanol a by P20 micropipette:

$$\frac{0.1}{5} \times 100\% = 2\%$$

The total percentage uncertainty is obtained from the sum of percentage uncertainty of ethanol concentrations, volume of *S. epidermidis*, and volume of ethanol.

Example (40% ethanol):

$$2.17\% + 0.5\% + 2\% = 4.67\%$$

Same calculations were made for other concentrations, results presented in Table 9.

Ethanol concentration (%)	% uncertainty for volume of 100% ethanol (%) (2 d.p)	% uncertainty for volume of autoclaved water (%) (2 d.p)	Total % uncertainty for ethanol concentration (%) (2 d.p)
40	0.50	1.67	2.17
45	2.22	1.82	4.04
50	2.00	2.00	4.00
55	1.82	2.22	4.04
60	1.67	0.50	2.17
65	1.54	0.57	2.11
70	1.43	0.67	2.10
75	1.33	0.80	2.13
80	1.25	1.00	2.25
85	1.18	1.33	2.51
90	1.11	2.00	3.11

Table 8. Percentage uncertainties for different ethanol concentrations.

Ethanol concentration (%)	% uncertainty for ethanol concentration (%) (2 d.p)	% uncertainty for volume of <i>S.epidermidis</i> (%)	% uncertainty for volume of ethanol (%)	Total % uncertainty (%) (2 d.p)
40	2.17	0.50	2.00	4.67
45	4.04	0.50	2.00	6.54
50	4.00	0.50	2.00	6.50
55	4.04	0.50	2.00	6.54
60	2.17	0.50	2.00	4.67
65	2.11	0.50	2.00	4.61
70	2.10	0.50	2.00	4.60
75	2.13	0.50	2.00	4.63
80	2.25	0.50	2.00	4.75
85	2.51	0.50	2.00	5.01
90	3.11	0.50	2.00	5.61

Table 9. Total percentage uncertainty in this investigation.

5 Analysis & Evaluation

5.1 Analysis of Data and Graph

40% to 55%

At ethanol concentrations between 40% to 55%, there were no measurable zones of inhibition after 24 or 48 hours of incubation, suggesting limited bactericidal effects. Hence, the hypothesis that there are significantl smaller zones of in-

hibitions below 50% is accepted. Due to *S. epidermidis*'s thick cell wall of peptidoglycan (15-80 nanometers) (Todar, 2015), low ethanol concentrations are insufficien to disrupt the peptidoglycan, hence ethanol molecules' inability to reach the membrane.

While the data for 40% to 50% ethanol aligns with McDonnell's study (McDonnell & Russell, 1999), 55% ethanol differs due to the different bacteria

used. McDonnell used *Escherichia coli*, a gram-negative bacteria with a thin peptidoglycan that is more susceptible to ethanol (UCSB ScienceLine, 2014), hence is killed with a lower concentration like 55% ethanol. This is supported by Morton's study, where gram-negative bacteria is killed by all ethanol concentrations from 40% to 100%, while gram-positive bacteria, as if the bacteria in this study, require 60% to 95% (Morton, 1950).

60% to 90%

At ethanol concentrations between 60% to 90%, the zone of inhibition indicated bactericidal effects. However, its effectiveness differed between different concentrations and incubation durations.

Figure 8 shows that after 24 hours, ethanol concentrations (x) and zone of inhibition (y) follow the inverse tangent function: $y = 0.96 \arctan(x - 60) + 1.52$. The r^2 value of 0.969 shows a strong correlation between the data and the trend-line, hence the accuracy of using inverse tangent to represent the data. The trend-line demonstrates an increase in the zone of inhibition between 55% to 65% ethanol, explained by the increased moles of ethanol performing bactericidal effects (Ingram, 1989). Based on the mathematical function, the optimum is between 90% to 100% with a zone of inhibition of 3.00 mm. However, the trend-line plateaus after 65%, and the two-sample t -test showed no significant differences in the zone of inhibition between 65% to 90%. Therefore, ethanol concentration between 65% to 90% had a similar level of bactericidal effects after 24 hours of incubation.

Figure 9 shows that after 48 hours, the two variables follow the inverse tangent function: $y = 0.8 \tan^{-1}(0.9x - 55) + 1.3$, where the trend-line plateaus after 70%. The r^2 value of 0.888 indicates a strong correlation between the data and the trend-line, but is weaker compared to Figure 8 due to the inconsistent decrease in the zone of inhibition between the two incubation periods. Based on the mathematical function, the optimum is between 90% to 100% with a zone of inhibition of 2.54mm.

Despite the trend-line showing an increase in the zone of inhibition between 55% to 70% ethanol, the two-sample t -tests showed no significant differences between 60% to 90%. Therefore, ethanol concentration between 60% to 90% has the same bactericidal effects after 48 hours of incubation.

Consequently, it rejects the hypothesis that the zone of inhibition decreases after 90%. This is due to evaporated ethanol molecules (vapor phase) being trapped in the closed agar plate, leading to collisions with liquid ethanol molecules on the agar surface, thus condensing to the liquid phase (Brown *et al.*, 2020). Therefore, high ethanol concentrations have enough treatment time with *S. epidermidis* for bactericidal effects in agar plates.

Differences between two incubation durations

The trend-line in Figure 9 demonstrates a downward vertical translation of 0.22 units from the trend-line in Figure 8, suggesting a 0.22 mm decrease in the zone of inhibition for all concentrations between 24 and 48 hours of incubation. However, most concentrations showed an inconsistent decrease. 65% ethanol showed a greater decrease of 1.1 mm; 75% ethanol showed a smaller decrease of 0.2 mm; 90% ethanol showed no changes. The paired t -tests indicated that 65% and 70% ethanol showed significant differences in the zone of inhibition between 24 and 48 hours of incubation, but 60%, 75%, 80%, 85%, and 90% showed no significant differences. Therefore, only 65% and 70% ethanol showed reduced bactericidal effects over time.

In real life, the Food and Drug Administration recommends an ethanol concentration of 60% to 95% in hand sanitizers (Antonucci, 2019). However, this investigation demonstrated a varying degree of bactericidal effects within the range. 60% ethanol had lower bactericidal effects compared to other concentrations; the bactericidal effects of 65% and 70% ethanol reduced over time. Therefore, individuals should choose hand sanitizer at ethanol concentrations between 70% to 95%, and should not solely rely on hand sanitizers in the longer period.

5.2 Evaluation of Data

All data for 40%, 45%, and 50% falls on the trend-line, has no standard deviation, and aligns with McDonnell's study, quoting "antimicrobial activity of ethanol is significantly lower at a concentration below 50%" [8], demonstrating high accuracy and reliability.

However, all data for 55% to 90% (except 65%, 75%, 80% in Figure 9) does not fall on the trend-line or align with other studies (McDonnell & Russell, 1999; Yamashita & Takeno, 2001), demonstrating low reliability. The large error bars for all data between 60% to 90% indicate a large spread of data between individual trials, potentially leading to a different trend-line. For example, after 24 hours of incubation, 65% and 90% ethanol only showed a 0.40 mm difference in the zone of inhibition, but the standard deviations were 1.33 mm and 0.98 mm respectively, significantly larger compared to the 0.4 mm difference. If 90% ethanol had a zone of inhibition of 2.22 mm (lower bound of the error bar), the two variables no longer exhibited an inverse tangent relationship, but instead, the trend predicted in the hypothesis. Therefore, due to the high standard deviations, the relationship between the two variables remains unclear after 50%.

The total percentage uncertainties for each ethanol concentration are low, ranging from 4.60% to 6.54%. The correlation coefficient between percentage uncertainty and the standard deviation is -0.638, suggesting a moderate negative correlation between the two variables. Therefore, the high standard deviations were caused by low sample size instead of the apparatus.

There was one anomalous data in this investigation. While all the zones of inhibition decreased or remained between 24 and 48 hours of incubation, 90% trial 1 showed a 1 mm increase. However, the paired *t*-test indicated that 90% ethanol showed no significant differences in the zone of inhibition between the two incubation durations, hence the anomalous data was inconsequential.

5.3 Evaluation of Sources

Scientific information on this report originated from a spectrum of reputable peer-reviewed journals and books, ensuring the validity of this research. The journals presented a detailed overview of their experiments with explanations on ethanol's antimicrobial mechanisms, allowing a holistic understanding of this research topic. However, they utilized different methodologies from this investigation, which hindered the comparison between the experimental and published data. To overcome this, further research was conducted to examine how the differences in data were influenced by different methodologies.

5.4 Evaluation of Methodology & Improvements

Strengths

All the control variables were controlled, such as using a P200 pipette to ensure 200 microlitres of *S. epidermidis* was added to each plate. Supported by the low total percentage uncertainties, this allowed an accurate basis of comparison.

An aseptic environment was maintained throughout the experiment to prevent external sources of contamination. The agar plates show no observed differences in bacterial colony morphology, hence no contamination of *S. epidermidis* (Antonucci, 2019).

Limitations of the disc diffusion method to model antimicrobial mechanisms on human skin

The disc diffusion method was safe and appropriate in a school laboratory, but presented many limitations for modeling the antimicrobial mechanisms of ethanol on human skin:

First, as ethanol concentration decreased when it diffused away from the paper disc on the agar surface, *S. epidermidis* within the zone of inhibition was killed by inconsistent concentrations, demonstrated by the imperfect circles around the paper discs. Hence, despite controlling the ethanol con-

centration added to the paper disc, the absolute concentration that inhibited *S. epidermidis* was unclear and uncontrolled, leading to large standard deviations and low accuracies of data.

Second, in low ethanol concentrations, the added water content slows down its evaporation rate on the human skin, allowing ethanol to penetrate through the microbial membranes. However, high ethanol concentrations also had enough treatment time with *S. epidermidis* for bactericidal effects in agar plates, hence the 90% ethanol demonstrated the same level of bactericidal effect as the literature optimal concentrations of 60% to 85% (McDonnell & Russell, 1999).

Third, the disc diffusion method relied on the diffusion of ethanol, which did not spread across the entire agar surface. Therefore, most *S. epidermidis* were not killed, shown by the observed turbidity after incubation. This led to significant bacterial growth in lower ethanol concentrations, for instance, the 1.1 mm decrease in the zone of inhibition with 65% ethanol. However, in real life, rubbing hands allow hand-sanitizers to cover all skin surfaces, thereby killing most microorganisms (Brown *et al.*, 2020), greatly inhibiting microbial growth.

Fourth, ethanol's direction of diffusion was unclear. It could diffuse into the agar instead of remaining on the surface to perform bactericidal effects.

These limitations can be avoided by the broth dilution method (Balouiri *et al.*, 2016). Culture tubes, containing a liquid medium and ethanol in desired concentration, are inoculated with *S. epidermidis*. Vortexing the tubes allows all *S. epidermidis* to be exposed to the same ethanol concentrations. After incubation, the tube's change in turbidity is examined by a UV-visible spectrometer (Balch, 1931) rather than observation, while bacterial growth is measured by a hemocytometer (Boundless, n.d.). Hemocytometer involves creating a volumetric grid divided into differently sized cubes to count the number of bacterial colonies, thereby accurately calculating the bacterial concentration.

Lastly, the characteristics of the agar medium and the human skin were different. The human skin has a slightly acidic pH of 4.7 to 5.75 (Eucerin, n.d.), which gives protection against microorganisms (cartron *et al.*, 2014), hence the data is not accurate under the human skin context. To improve, pH 5 can be added to the liquid medium mentioned above to simulate the pH of human skin.

Other limitations

Despite ethanol's same mechanism on the bacterial and viral membrane, there was doubt whether the specific concentrations were efficacious for SARS-CoV-2. As viruses do not have cell walls, bacteria that lack a cell wall, including *Mycoplasma* and L-form bacteria (Boundless, n.d.) could be a more appropriate model for SARS-CoV-2.

6 trials were conducted for each ethanol concentration to reduce the standard deviation, but the high values suggest an insufficient number of trials conducted. For example, the zone of inhibition for 65% ethanol ranged from 1 mm to 5 mm (Table 4), hence it was difficult to identify whether the values were flu es or normal cases. To improve, the number of trials is increased to 12 trials to further reduce standard deviation.

There were only two changes in the incubation duration: 24 and 48 hours, hence unclear whether ethanol concentrations between 75% to 90% continue to have no significant differences in the zone of inhibition beyond 48 hours. The large time interval also made it unclear whether 40% and 55% ethanol showed no bactericidal effects, or *S. epidermidis* grew back into the zone of inhibition within 24 hours. To improve, there should be at least five changes in the incubation durations. The reproduction rate of *S. epidermidis* should also be examined in the preliminary investigation to make an appropriate time interval adjustment.

6 Conclusion

6.1 Conclusion

Ethanol concentrations between 40% to 55% showed no zone of inhibition, suggesting limited bactericidal effects. Ethanol concentration between 60% to 90% showed zone of inhibitions after 24 and 48 hours of incubation, hence bactericidal effects in the short and long term. Specifically, concentrations between 65% to 90% showed the same level of bactericidal effect in the short term, but 75% had greater inhibition ability in the long term.

6.2 Further Investigation

Studies have shown that most alcohol-based hand antiseptics contain 1-propanol, 2-propanol, or ethanol (Brown *et al.*, 2020). Ethanol contains two-carbons, while both 1-propanol and 2-propanol contain three-carbons but differ by the position of the hydroxyl group (Figure 10). The different chemical structures suggest different levels of antimicrobial effects. Therefore, the antimicrobial effects of 1-propanol and 2-propanol can be explored as a further investigation to evaluate the most effective agent for microbial inhibition.

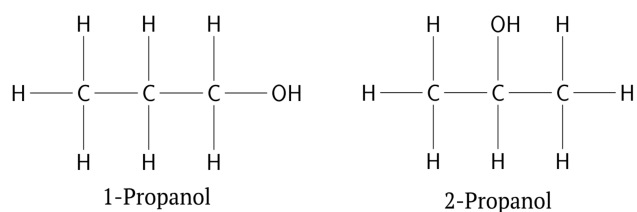


Fig. 10. Comparison between propan-1-ol and propan-2-ol.

References

- Antonucci, S. (2019). *How to Detect Contamination in Agar Plates.* <https://careertrend.com/how-6101553-detect-contamination-agar-plates.html>
- Balch, R. T. (1931). Measurement of Turbidity with a Spectrophotometer. *Industrial Engineering Chemistry Analytical Edition*, 3, 124-127. doi: 10.1021/ac50074a003
- Balouiri, M., Sadiki, M., & Ibensouda, S. K. (2016). Methods for in vitro evaluating antimicrobial activity: A review. *Journal of Pharmaceutical Analysis*, 6, 71-79. doi: 10.1016/j.jpha.2015.11.005
- Black, R. (2010). *Hand sanitizer only last for two minutes, not effective at killing germs long-term: research.* <https://www.nydailynews.com/life-style/health/hand-sanitizer-minutes-not-effective-killing-germs-long-term-research-article-1.190969>
- Boundless. (n.d.). *Boundless Microbiology.* [https://bio.libretexts.org/Bookshelves/Microbiology/Book:_Microbiology_\(Boundless\)](https://bio.libretexts.org/Bookshelves/Microbiology/Book:_Microbiology_(Boundless))
- Brown, LeMay, Busten, Murphy, Woodward, & Halpern, J. (2020). 11.5: Vapor Pressure. [https://chem.libretexts.org/Bookshelves/General_Chemistry/Map:_Chemistry_-_The_Central_Science_\(Brown_et_al.\)/11:_Liquids_and_Intermolecular_Forces/11.5:_Vapor_Pressure](https://chem.libretexts.org/Bookshelves/General_Chemistry/Map:_Chemistry_-_The_Central_Science_(Brown_et_al.)/11:_Liquids_and_Intermolecular_Forces/11.5:_Vapor_Pressure)
- Bruslind, L. (2020). Bacteria: Cell Walls. Oregon State University. <https://open.oregonstate.edu/education/generalmicrobiology/chapter/bacteria-cell-walls/>
- Cao, H., Wei, D., Yang, Y., Shang, Y., Li, G., Zhou, Y., ... Xu, Y. (2017). Systems-level understanding of ethanol-induced stresses and adaptation in *E. coli*. *Scientific Reports*, 7. doi: 10.1038/srep44150
- Cartron, M. L., England, S. R., Chiriac, A. I., Josten, M., Turner, R., Rauter, Y., ... Foster, S. J. (2014). Bactericidal Activity of the Human Skin Fatty Acid cis-6-Hexadecanoic Acid on *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 58, 3599-3609. <https://aac.asm.org/content/58/7/3599> doi: 10

- .1128/AAC.01043-13
- Centers for Disease Control and Prevention. (2019). *Hand Hygiene in Healthcare Settings*. <https://www.cdc.gov/handhygiene/index.html>
- Eucerin. (n.d.). *Understanding skin – Skin’s pH*. <https://int.eucerin.com/about-skin/basic-skin-knowledge/skins-ph>
- French Agency for Food, Environmental and Occupational Health Safety. (2016). *Assessing the Risks of Ethanol*. <https://www.anses.fr/en/content/assessing-risks-ethanol>
- Golin, A. P., Choi, D., & Ghahary, A. (2020). Hand Sanitizers: A Review of Ingredients, Mechanisms of Action, Modes of Delivery, and Efficacy Against Coronaviruses. *American Journal of Infection Control*, 48, 1062-1067. <https://www.sciencedirect.com/science/article/pii/S0196655320305629> doi: 10.1016/j.ajic.2020.06.182
- Huffer, S., Clark, M. E., Ning, J. C., Blanch, H. W., & Clark, D. S. (2011). Role of Alcohols in Growth, Lipid Composition, and Membrane Fluidity of Yeasts, Bacteria, and Archaea. *Applied and Environmental Microbiology*, 77, 6400–6408. <https://aem.asm.org/content/77/18/6400> doi: 10.1128/AEM.00694-11
- Huynh-Delerme, C., Artigou, C., Bodin, L., Tardif, R., Charest-Tardif, G., Verdier, C., ... Desmares, C. (2012). Short Communication: Is Ethanol-Based Hand Sanitizer Involved in Acute Pancreatitis after Excessive Disinfection?—An Evaluation with the Use of PBPK Model. *Journal of Toxicology*, 2012, 1-7. doi: 10.1155/2012/959070
- Ingram, L. O. (1989). Ethanol Tolerance in Bacteria. *Critical Reviews in Biotechnology*, 9, 305-319. doi: 10.3109/07388558909036741
- Ingólfsson, H. I., & Andersen, O. S. (2011). Alcohol’s Effects on Lipid Bilayer Properties. *Biophysical Journal*, 101, 847-855. doi: 10.1016/j.bpj.2011.07.013
- Jefferson, T., Spencer, E., Plüddemann, A., Roberts, N., & Heneghan, C. (2020). *Transmission Dynamics of COVID-19*. <https://www.cebm.net/evidence-synthesis/transmission-dynamics-of-covid-19/>
- Karki, G. (2017). *Bacterial Cell wall: Structure, Composition and Types*. <https://www.onlinebiologynotes.com/bacterial-cell-wall-structure-composition-types/>
- Laerd Statistics. (2018). *One-way ANOVA*. <https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide-2.php>
- McDonnell, G., & Russell, A. D. (1999). Antiseptics and Disinfectants: Activity, Action, and Resistance. *Clinical Microbiology Reviews*, 12, 147-179. <https://cmr.asm.org/content/12/1/147> doi: 10.1128/cmr.12.1.147
- Microbiology in Schools Advisory Committee. (2010). *Safety guidelines*. http://www.ncbe.reading.ac.uk/SAFETY/PDF/MISAC_safety_guidelines.pdf
- Microbiology Society. (2020). *Routes of transmission | Microbes and the human body*. <https://microbiologysociety.org/why-microbiology-matters/what-is-microbiology/microbes-and-the-human-body/routes-of-transmission.html>
- Morton, H. E. (1950). THE RELATIONSHIP OF CONCENTRATION AND GERMICIDAL EFFICIENCY OF ETHYL ALCOHOL. *Annals of the New York Academy of Sciences*, 53, 191-196. doi: 10.1111/j.1749-6632.1950.tb31944.x
- Namvar, A. E., Bastarahang, S., Abbasi, N., Ghehi, G. S., Farhadbakhtiaran, S., Arezi, P., ... Chermahin, S. G. (2014). Clinical characteristics of Staphylococcus epidermidis: a systematic review. *GMS hygiene and infection control*, 9(3). doi: 10.3205/dgkh000243
- Nedwell, D. B. (1999). Effect of low temperature on microbial growth: lowered affinity for substrates limits growth at low temperature. *FEMS Microbiology Ecology*, 30, 101-111. <https://academic.oup.com/femsec/article/30/2/101/608571> doi: 10.1111/j.1574-6941.1999.tb00639.x
- Ophardt, C. E. (2003). *Denaturation Protein*. Elmhurst College. <http://chemistry.elmhurst.edu/vchembook/568denaturation.html#:~:text=Alcohol%20denatures%20proteins%20by%20disrupting>
- Osman, Y. A., & Ingram, L. O. (1985). Mechanism of ethanol inhibition of fermentation in *Zymomonas mobilis* CP4. *Journal of Bacteriology*, 164, 173-180. <https://jb.asm.org/>

content/jb/164/1/173.full.pdf doi: 10.1128/
jb.164.1.173-180.1985

Todar, K. (2015). *Online Textbook of Bacteriology*. University of Wisconsin. http://www.textbookofbacteriology.net/structure_5.html#:~:text=The%20outer%20membrane%20is%20attached,of%20several%20layers%20of%20peptidoglycan.

UCSB ScienceLine. (2014). *How does hand sanitizer kill germs?* <http://scienceline.ucsb.edu/getkey.php?key=4374#:~:text=Alcohols%20kill%20germs%20by%20destroying>

van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., ... Munster, V. J. (2020). Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *New England Journal of Medicine*, 382, 1564-1567. doi: 10.1056/nejmc2004973

Wiencek, M. (2020). *Are All Viruses the Same? Nope!* <https://www.contecinc.com/articles/viruses-are-not-the-same/>

World Health Organization. (2009). *Who guidelines on hand hygiene in health care: a summary*. https://www.who.int/gpsc/information_centre/hand-hygiene-summary/en/

Worldometer. (2020). *Coronavirus toll update: Cases deaths by country*. <https://www.worldometers.info/coronavirus/>

Yamashita, M., & Takeno, A. (2001). Relationship Between Bactericidal Activity and the Hydrophobicity-Hydrophilicity Balance of Alcohol Solutions. *Biocontrol Science*, 6, 107-111. https://www.jstage.jst.go.jp/article/bio1996/6/2/6_2_107/_pdf doi: 10.4265/bio.6.107

Zaharia, D. C., Iancu, C., Steriade, A. T., Muntean, A. A., Balint, O., Popa, V. T., ... Bogdan, M. A. (2010). MicroDSC study of Staphylococcus epidermidis growth. *BMC Microbiology*, 10, 322. doi: 10.1186/1471-2180-10-322

“Statistics conceal as much as they reveal.”

Denton Philtjens 費丹誠

“The US COVID-19 death-toll has passed 500,000” (Feuer), is a recent statistic which reveals knowledge about the trajectory of the pandemic. However, it also doesn’t inform us about the virus’ survival rate or the impact of pre-existing health conditions, which can lead the statistic’s audience to compose false conclusions. Hence, it is evident that statistics, which are tools that relate to the collection and analysis of data, can have both a revealing and concealing role during knowledge production. In this context, revealing refers to facilitating knowledge production by divulging beneficial and truthful knowledge, whilst concealing refers to hindering knowledge production by hiding reality and thus prompting the production of false knowledge. Statistics can broadly be categorised into inferential and descriptive, and they are used in the human sciences (HS) and natural sciences (NS). Responding to the claim “statistics conceal as much as they reveal”, I will argue that the level of concealment and revealment differs depending on the AOK (for they are different by nature), and the category of statistical analysis. Moreover, I will explore how inferential statistics are more revealing in NS whilst descriptive statistics are equally revealing in both AOKs, and that usually statistics reveal more than they conceal in both AOKs, although these arguments don’t always hold true.

Firstly, inferential statistics are tools which analyze and help us understand bivariate data, facilitating the discovery of previously unknown patterns (knowledge). Typically, they are more revealing in NS due to a greater ability to isolate and manipulate variables. For instance, when I conducted my Physics Internal Assessment investigating the relationship between salinity and specific heat capacity, I controlled most of my variables and incrementally increased the salinity of

my sample. Then, through taking measurements with a thermometer, I collated data which later underwent linear regression analysis to find the optimal trendline. Here, statistical analysis allowed me to produce knowledge (a trend) with relative accuracy due to the controlled environment, revealing information about our natural world. On the contrary, HS scientists don’t often have this ability to manipulate and isolate variables. For instance, in Economics, when Neumark and Wascher conducted research on the minimum wage, they observed historical data and compared many US states and the impact of minimum wages on employment growth (Matthews). Hence, they didn’t manually manipulate the independent variables and had to make broad assumptions about the populations studied. This was done, because HS researchers have variables that involve human subjects and large demographics, which can be harmful and logistically difficult to manually manipulate and control. This can produce reliability issues with the findings, and often a third variable problem can arise, which is where the two observed variables don’t cause each other to change and an underlying variable is responsible. This problem has arisen in the minimum wage context, as Neumark and Wascher initially found that in the 1980s, minimum wages in blue states substantially decreased employment growth compared to red states without a minimum wage. However, it was later revealed that because blue states experienced an economic downturn that predated a minimum wage hike, in reality blue states had lower employment growth and separately increased their minimum wages (Matthews). Hence, Neumark and Wascher’s statistical analysis led their report to produce a misleading conclusion (based on our current understanding), concealing our at-the-time understanding of reality. And whilst this concealment is currently corrected, it has still

produced detrimental and long-lasting effects, for statistics are powerful as they appeal to reason, and in this case were used in the justification for US government public policy.

However, whilst I previously argued that NS scientists have a greater ability to control and manipulate variables than HS scientists, this is sometimes not the case, especially when it comes to large-scale or human related investigations, where there can be logistical difficulties and ethical dilemmas. For instance, when the IPCC creates climate change predictions, they rely on extrapolations based on historical data, for they can't logistically conduct large-scale practical experiments. This leads their analysis to incorporate uncontrolled variables and assumptions, for instance regarding the degree to which climate change engages in a positive feedback loop (Carter), which causes certain people to question how revealing the knowledge claims produced are. Additionally, whilst I previously argued that inferential statistics are generally revealing in both AOKs, this is false when the equipment used produces undetected systematic errors. For instance, in the 1970s, Joseph Weber claimed that he identified gravitational waves using a proprietary method. However, it was later discovered that his equipment was systematically flawed, and that the waves he had detected were from other interferences (Koberlein). Whilst we currently have proof of the existence of gravitational waves, it has been proven that Weber's method simply couldn't have detected gravitational waves, and that his conclusion resulted from a disguised error. Hence, the false knowledge which he produced through statistical analysis was debunked through an in-depth analysis of his methodology and raw data by other scientists. Therefore, in the long-run, false knowledge which conceals reality will be corrected when contrary evidence is proposed. Lastly, whilst I previously argued that HS statistics which lead to false conclusions can have detrimental practical implications, there are often no better alternatives. In fact, because statistics allow us to transform large-scale knowledge produced via sense perception and measurements into methodically organized reasoning, they play a vital role in both

AOKs. Moreover, beyond simply increasing the volume of our knowledge, these statistics also increase the precision of our knowledge, as numerical descriptions are far more precise than linguistic descriptions such as "large decrease". Hence, the creation of misleading information is arguably a necessary evil, and the concealment which inferential statistics can cause is simply part of the scientific process, where knowledge slowly gets refined through peer-review. Hence, inferential statistics are more revealing than concealing in the long-run.

Moving on, descriptive statistics are tools such as central tendencies (mean, median and mode) and standard deviation, which analyze various features of a univariate data set. In both AOKs, these tools are equally effective at revealing reality by summarizing the characteristics of data sets and measuring uncertainty, for the nature of the two AOKs doesn't impact these abilities. For instance, these statistical techniques are employed in the NIST-IUPAC database for solubility, which describes the solubility of various chemicals under given conditions. Firstly, by calculating the mean solubility across various research papers and trials, knowledge regarding the "true" solubility value under given conditions can be revealed. Moreover, measuring the standard deviation between different research papers can help reveal how trustworthy and established the knowledge produced is, by seeing whether there is a scientific consensus. For instance, in the NIST-IUPAC database for SO₂'s solubility, across various research papers there is a standard deviation of 4.6%, which indicates that there is further room for research and more advanced equipment (Young). Additionally, in certain cases this can also signify that paradigm-shifting "revolutionary science" breakthroughs might be evident.

However, descriptive statistics can be more concealing than revealing in both AOKs when measures of central tendency are used to explain very broad sets of data with uncontrolled underlying variables. For example, the NS literature on pleural Mesothelioma (a cancer) suggests that the median mortality after diagnosis is 6-12 months (Molinari). Whilst many patients might interpret

this statistic at face value, and lose hope about their survival, this is an inadequate conclusion, for an average value doesn't indicate (and hence conceals) the fact that there are a range of mortality rates, and that many variables can influence the patient's mortality, such as age and genetics. Hence, the third variable problem once again arises, and this statistic hinders the communication of knowledge, for the general public might be unfamiliar with the limitations of statistics and might interpret means and medians as set realities, thus causing their understanding of "the full picture" to be concealed, prompting them to produce false conclusions. Moreover, factual statistics which might prompt particular interpretations of a topic can be misused to portray specific views. This is prevalent in HS, for instance regarding China's economic status. Whilst the average monthly income of 1070 USD (CEIC Data) and the monthly median income of 520 USD (Cyrill) are both factually correct, the average doesn't factor inequality, the negligence of which politicians might abuse to propel certain perceptions of the nation. Hence, statistical literacy and not drawing too many implications from descriptive statistics is key to limiting any concealment. Moreover, it must also be acknowledged that overall, in both AOKs, descriptive statistics are usually more revealing than concealing, because they produce knowledge and indicate areas for improvements, two roles which are otherwise difficult to achieve in a large-scale methodical manner.

Hence, it is evident that while both descriptive and inferential statistics often have a revealing role during knowledge production, the degree to which can differ due to the AOK. However, circumstances involving broad data sets and flawed equipment can cause statistics to be more concealing than revealing. Nonetheless, through the careful evaluations of the raw data and research methodologies, a lot of these concealments are eradicated. However, if the false knowledge created based on misleading statistics is applied to public policy, even if any concealments are eventually "corrected" academically, the short-term impact can still be detrimental and long-lasting. This exploration is certainly meaningful, for statistics are used heavily in

research and therefore understanding the concealing and revealing nature of statistics is important to assess the reliability and utility of the claims created. Moreover, living in the digital age with access to wide-ranging sources and information, it is more important than ever to understand the implications and nature of statistics in order to make informed decisions in our daily lives.

Works Cited

- Carter, Robert M. "The myth of dangerous human-caused climate change." Australasian Institute of Mining and Metallurgy. 2007. <https://researchonline.jcu.edu.au/3130/>.
- CEIC Data. "China Monthly Earnings, 2000 – 2021 Data." <https://www.ceicdata.com/en/indicator/china/monthly-earnings#:~:text=China%20Monthly%20Earnings%20stood%20at>.
- Cyrill, Melissa. "Who Make Up China's Middle Class? We Asked 5 Simple Questions." Feb. 2019. <https://www.china-briefing.com/news/chinas-middle-class-5-questions-answered>.
- Feuer, Will. "U.S. surpasses 500,000 Covid deaths after yearlong battle with pandemic." Feb. 2021. <https://www.msn.com/en-us/money/companies/u-s-surpasses-500-000-covid-deaths-after-yearlong-battle-with-pandemic/ar-BB1dUELK>.
- Koberlein, Brian. "Joseph Weber And The Failed Search For Gravitational Waves." Oct. 2015. <https://www.forbes.com/sites/briankoberlein/2015/10/12/joseph-weber-and-the-failed-search-for-gravitational-waves/?sh=671a07311593>.
- Matthews, Dylan. "Should the minimum wage be raised? The economic debate, explained." Nov. 2019. <https://www.vox.com/future-perfect/2019/11/20/20952151/should-minimum-wage-be-raised>.
- Molinari, Linda. "Mesothelioma Life Expectancy." Oct. 2020. <https://www.mesothelioma.com/mesothelioma/prognosis/life-expectancy/>.
- Young, Colin L. "SULFUR DIOXIDE, CHLORINE, FLUORINE AND CHLORINE OXIDES." *IUPAC-NIST Solubility Data Series 12* (1983). <https://srdata.nist.gov/solubility/IUPAC/SDS-12/SDS-12.pdf>.

Fourier Series modelling with applications to electrocardiogram readings

Dionne Daiyin Yeung 楊岱殷

Introduction

Electrocardiogram (ECG) readings are used to measure the beating of a heart, and provides pertinent information about the heart's current health and condition. I wondered if these readings could be modelled with mathematical functions, so they can serve as a mathematical premise to quantify and distinguish between hearts. Here, the Fourier series is used to model a healthy sinus rhythm.

1 Background

1.1 Overview of the Fourier Series

The Fourier series was developed by Joseph Fourier in the early 19th century when he was working with heat equations in physics. In which, a new solution to an equation is formed by the summation (in other words, the addition) of two trigonometric functions. He wondered if a square wave, or any function, can simply be expressed as the summation of sine or cosine, even if they look nothing like these trigonometric functions (see figure 1). The result, originally only applied for heat, is now used as the basis for a wide array of scientific works, from quantum mechanics to noise cancelling headphones (3Blue1Brown, 2019).

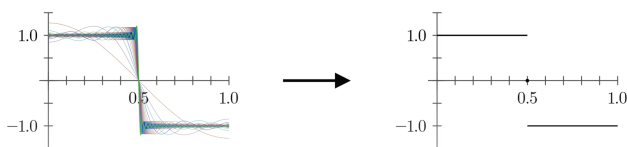


Fig. 1. Square wave formed by the Fourier series with increasing amounts of sine and cosine functions in different colors (left), with the original function (right) (3Blue1Brown, 2019).

The foundation of the Fourier series lies in trigonometry, particularly sine and cosine functions.

Both functions are periodic, in that they repeat at regular intervals with a period (T) (see figure 2).

1.2 Definition of the Fourier Series

The Fourier series has two forms: trigonometric and exponential. The math explained here is in trigonometric form (real numbers), rather than exponential (complex numbers), simply because this exploration deals with real values rather than complex. The Fourier series is a periodic function formed from the superposition of sine and cosine functions (see figure 3).

Specific functions can therefore be expressed as the sum of infinite sine and cosine functions, where each function becomes increasingly more frequent. Because deriving Fourier coefficients assumes that $T = 2\pi$, when approximating functions with a different period, another coefficient is required, so that the Fourier series can horizontally compress or stretch to match the original function. This can be done so through using the radial frequency as a coefficient. For a trigonometric function $f(x)$, the rate of angular change is known as the radial frequency (ω), and is hence $\omega = \frac{\Delta\theta}{\Delta x}$. As the change in angle, in radians, for each period is 2π , it can also be expressed as $\omega = \frac{2\pi}{T}$. In this context, ω can simply be understood as the period relative to 2π , and is hence a factor that horizontally compresses or stretches the Fourier series to the correct period.

Considering some function $f(x)$ with a radial frequency ω , the Fourier Series is expressed below.

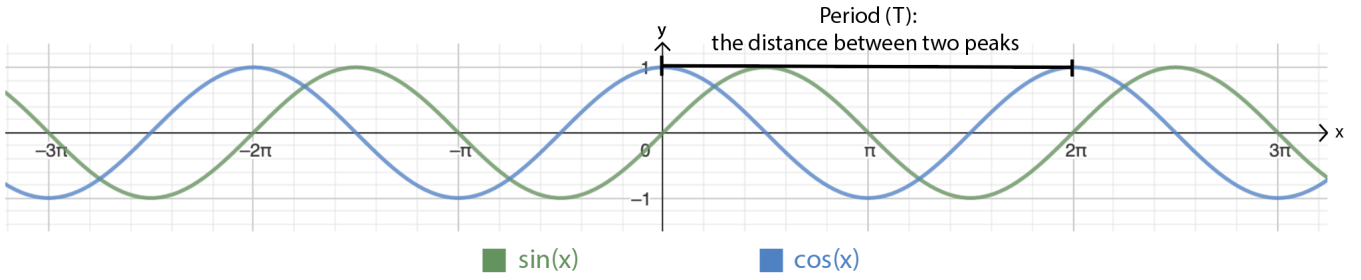


Fig. 2. Visualization of period, with sine and cosine waves graphed on GeoGebra (Hohenwarter *et al.*, 2013).

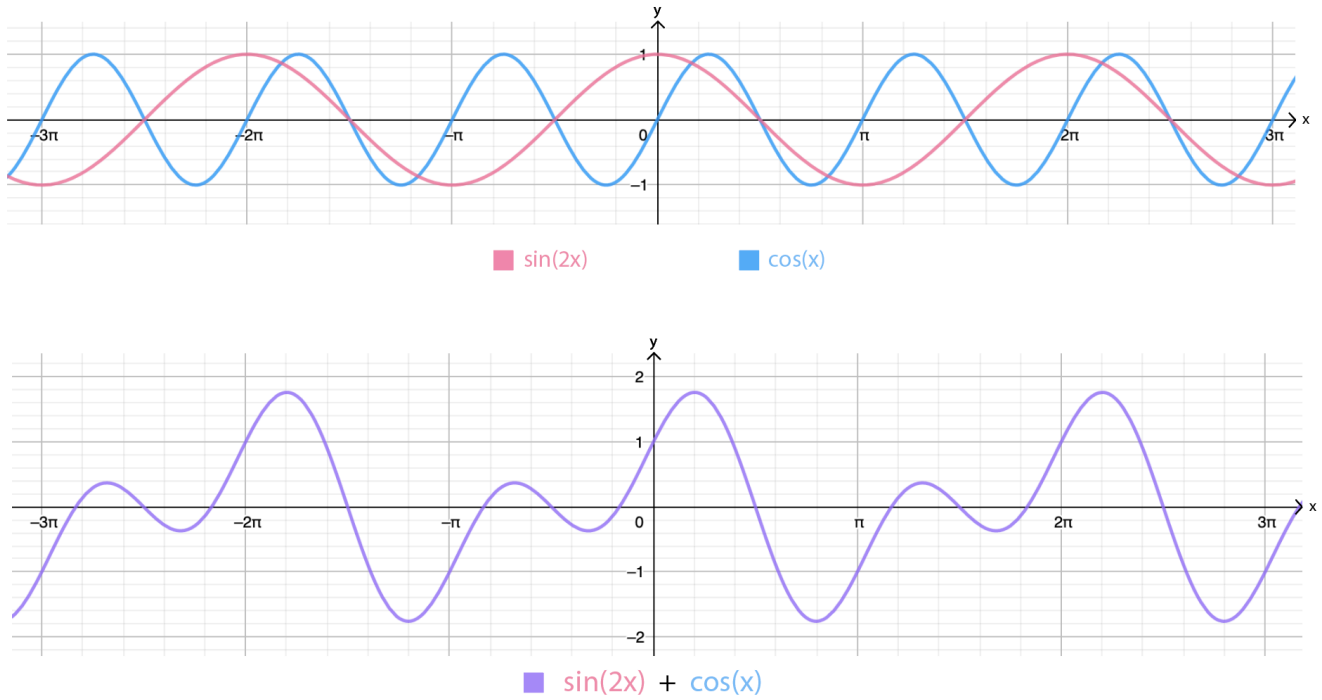


Fig. 3. Visualization of the superposition of trigonometric functions, where two functions (above) can be superposed to create a new function (below). Graphed on GeoGebra (Hohenwarter *et al.*, 2013).

Definition 1.1. When $n \in \mathbb{Z}^+$ and $a_n, b_n \in \mathbb{R}$ (Chou, n.d.):

$$f(x) = \sum_{n=0}^{\infty} (a_n \cos(\omega n x) + b_n \sin(\omega n x)) \quad (1)$$

Definition 1.2. When $n = 0$:

$$a_0 \cos(\omega(0)x) = a_0 \cos(0) = a_0$$

$$b_0 \sin(\omega(0)x) = b_0 \sin(0) = 0$$

Thus, the Fourier series can also be rewritten as:

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(\omega n x) + b_n \sin(\omega n x)) \quad (2)$$

The n constants signify the frequency of the trigonometric function, and through increasing n , the function gets more accurate to the modelled equation.

This increase in accuracy is demonstrated in figure 5. The coefficient a_0, a_n , and b_n determine the amplitude of each trigonometric function. The summation of different amplitudes would influence their superposition, thus determining what the final function would look like (see figure 4).

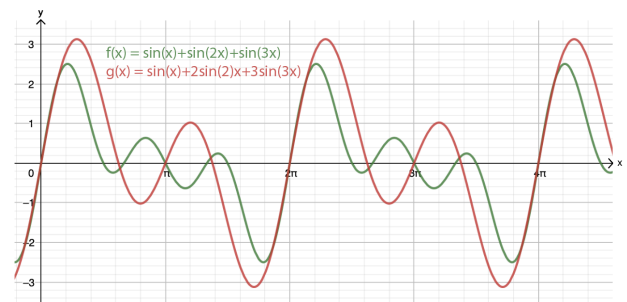


Fig. 4. Two functions of the Fourier series composed of 3 n values with differing b_n coefficients revealing how simple changes in amplitude, even in just one coefficient, lead to differing overall functions.

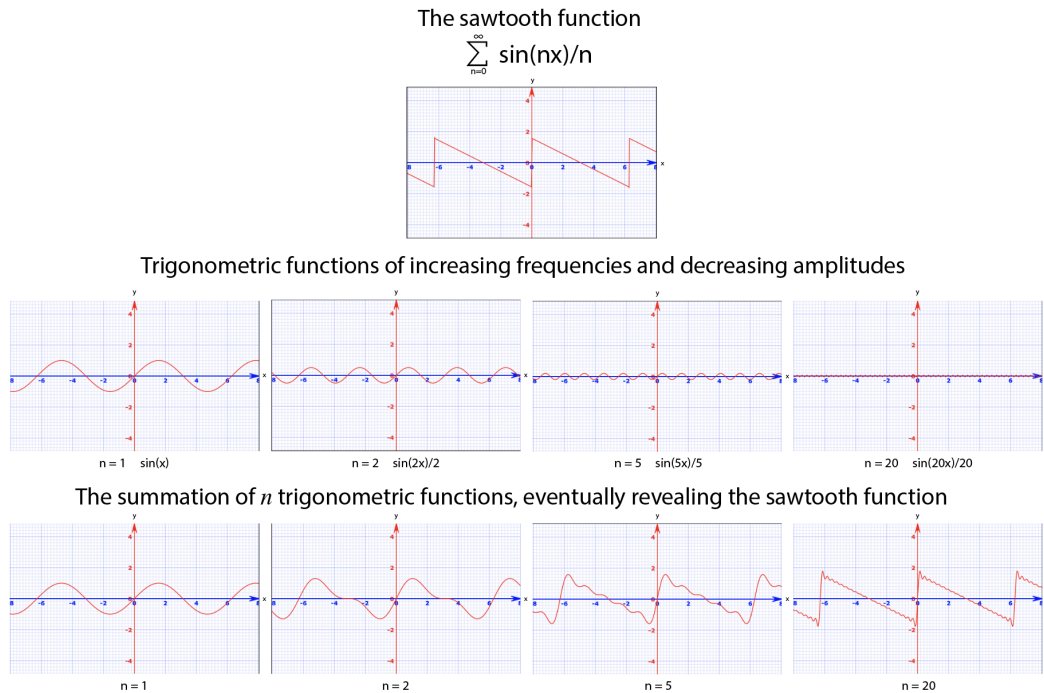


Fig. 5. To visualize just how the Fourier series works, this figure demonstrates how the summation of increasing values of n and changes in coefficient a_0 , a_n , and b_n slowly lead to a sawtooth function through the superposition of different trigonometric functions.

1.3 Electrocardiogram Readings

Electrocardiogram readings measure electric impulses from a beating heart. In hospitals, they are essential in monitoring heart health by the shape and rhythm of these electrical impulses. Through this, they can detect a myriad of heart conditions, such as a stroke, arrhythmia, etc. and immediately intercept it before any serious consequences occur.

Changes in amplitude, shape, and rhythm of these waves would all reflect the contractions and relaxations of heart chambers, and hence give crucial information about heart health. A simple change in the ECG, like a depression in the T wave, can indicate severe heart conditions such as ischemia - a shortage of oxygen supplied to the heart (Price, 2010).

With the Fourier series, the shapes of these waves would be determined by the a_0 , a_n , and b_n coefficients. The numeric value of these coefficients can therefore be crucial indicators of heart conditions and disease, and thereby differentiate between them. It may even be able to detect smaller changes that are not obvious to the naked eye, and catch heart issues before they happen.

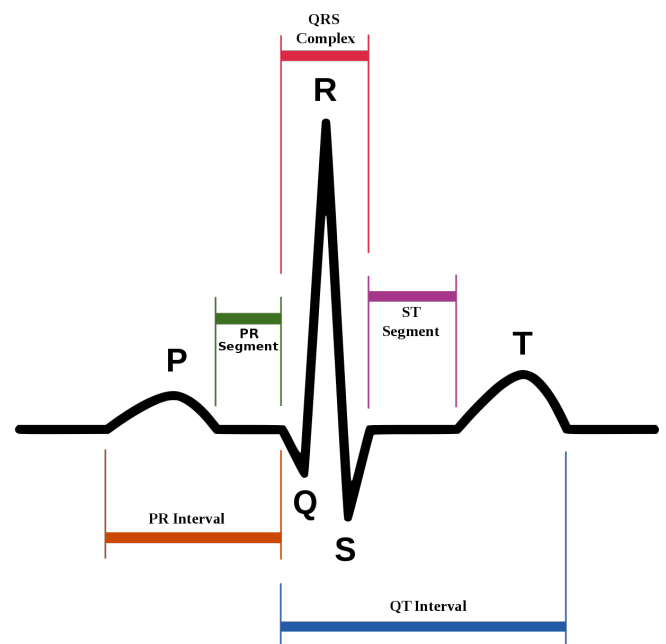


Fig. 6. "Schematic diagram of normal sinus rhythm for a human heart as seen on ECG (with English labels)" (Atkielski, 2007).

Further, the frequency of heart beats also plays a role in heart health. Mathematically, it is signified by the period. This is accounted for in the Fourier series through ω . Changes in the period can also indicate abnormalities in heart health (see figure 6 and 7).



Fig. 7. Unhealthy ECG caused by ischemia (decreased blood flow and oxygen to the heart muscle) (Anderson & DiCarlo, 2000).

However, it is most important to consider what a healthy ECG looks like first numerically, hence the purpose of this study.

2 Deriving Fourier Coefficient

The next step to modelling the ECG is to hence determine the coefficient a_0 , a_n , and b_n . To determine these coefficients a general equation for the coefficient must be derived, then applied to ECG readings.

2.1 Trigonometric Integration

Some integration proofs are first required, as they will be used to derive equations for finding Fourier coefficients

Proof 1: Integrals of Sine and Cosine

Claim 1: for $m \in \mathbb{Z}$:

$$\int_0^T \sin(\omega m x) \, dx = 0 \quad (3)$$

$$\int_0^T \cos(\omega m x) \, dx = 0 \quad (4)$$

The proof for both equations are similar, hence the proof is demonstrated for equation 3 as an example. By taking the anti-derivative of the chain rule:

$$\int_0^T \sin(\omega m x) \, dx = - \left[\frac{1}{\omega m} \cos(\omega m x) \right]_0^T$$

Substituting T and 0 into x :

$$\begin{aligned} \int_0^T \sin(\omega m x) \, dx &= - \frac{1}{\omega m} (\cos(\omega m T) - \cos(\omega m 0)) \\ &= - \frac{1}{\omega m} (\cos(\omega m T) - 1) \end{aligned}$$

Recalling that $\omega = \frac{2\pi}{T}$, we can substitute the variable:

$$\begin{aligned} \int_0^T \sin(\omega m x) \, dx &= - \frac{1}{\omega m} (\cos\left(\frac{2\pi}{T} m T\right) - 1) \\ &= - \frac{1}{\omega m} (\cos(2\pi m) - 1) \end{aligned}$$

As cosine is periodic, any integer multiple of 2π effectively adds a period onto another period, and is thus equivalent to simply 2π . Hence, the equation simplifies to:

$$\begin{aligned} \int_0^T \sin(\omega m x) \, dx &= - \frac{1}{\omega m} (\cos(2\pi) - 1) \\ &= - \frac{1}{\omega m} (1 - 1) = 0 \end{aligned}$$

Proof 2: Integrals of the Products of Sine and Cosine

Claim 2.1: for $m, n \in \mathbb{Z}^+$:

$$\int_0^T \sin(\omega n x) \cos(\omega m x) \, dx = 0 \quad (5)$$

The proof is as follows. By using the product to sum identity, the equation can be rewritten as:

$$\begin{aligned} &\int_0^T \sin(\omega n x) \cos(\omega m x) \, dx \\ &= \int_0^T \frac{1}{2} (\sin(\omega(m+n)x) + \sin(\omega(m-n)x)) \, dx \end{aligned}$$

Using the distributive property of integrals:

$$\begin{aligned} &\int_0^T \sin(\omega n x) \cos(\omega m x) \, dx \\ &= \frac{1}{2} \int_0^T \sin(\omega(m+n)x) \, dx + \frac{1}{2} \int_0^T \sin(\omega(m-n)x) \, dx \end{aligned}$$

As m and n are integers, $m+n$ and $m-n$ are also going to be integers. We can thus apply Claim 1, to prove that:

$$\int_0^T \sin(\omega n x) \cos(\omega m x) \, dx = \frac{1}{2} 0 + \frac{1}{2} 0 = 0$$

Claim 2.2: for $m, n \in \mathbb{Z}^+$:

$$\int_0^T \sin(\omega n x) \sin(\omega m x) \, dx = \begin{cases} 0 & \text{if } m \neq n \\ \frac{1}{2} T & \text{if } m = n \end{cases} \quad (6)$$

$$\int_0^T \cos(\omega n x) \cos(\omega m x) \, dx = \begin{cases} 0 & \text{if } m \neq n \\ \frac{1}{2} T & \text{if } m = n \end{cases} \quad (7)$$

As both proofs are quite similar, the proof for Equation 6 is demonstrated as an example. With the product to sum identity:

$$\begin{aligned} & \int_0^T \sin(\omega n x) \sin(\omega m x) \, dx \\ &= \int_0^T \frac{1}{2} (\cos(\omega(m-n)x) - \cos(\omega(m+n)x)) \, dx \\ &= \frac{1}{2} \int_0^T \cos(\omega(m-n)x) \, dx - \frac{1}{2} \int_0^T \cos(\omega(m+n)x) \, dx \end{aligned}$$

For case $m \neq n$:

As $m+n$ and $m-n$ are both non-zero integers, by applying Claim 1, we get:

$$\begin{aligned} & \int_0^T \sin(\omega n x) \sin(\omega m x) \, dx \\ &= \frac{1}{2} \int_0^T \cos(\omega(m-n)x) \, dx - \frac{1}{2} \int_0^T \cos(\omega(m+n)x) \, dx \\ &= \frac{1}{2} \cdot 0 - \frac{1}{2} \cdot 0 = 0 \end{aligned}$$

For case $m = n$:

$$\begin{aligned} & \int_0^T \sin(\omega n x) \sin(\omega m x) \, dx \\ &= \frac{1}{2} \int_0^T \cos(\omega(m-n)x) \, dx - \frac{1}{2} \int_0^T \cos(\omega(m+n)x) \, dx \\ &= \frac{1}{2} \int_0^T \cos(\omega(0)x) \, dx - \frac{1}{2} \int_0^T \cos(\omega(2m)x) \, dx \\ &= \frac{1}{2} \int_0^T 1 \, dx - \frac{1}{2} \int_0^T \cos(\omega(m+n)x) \, dx \end{aligned}$$

By applying Claim 1:

$$\begin{aligned} \int_0^T \sin(\omega n x) \sin(\omega m x) \, dx &= \frac{1}{2} \int_0^T 1 \, dx - 0 \\ &= \left[\frac{1}{2} x \right]_0^T \\ &= \frac{1}{2} (T - 0) = \frac{1}{2} T \end{aligned}$$

Each claim has been successfully proven, and it will ease the derivation process.

2.2 Formal Derivations

In light of these proofs, Fourier's coefficient can now be derived to mathematically model ECG readings (Khan, 2008; Cheever, 2019).

Coefficient a_n

To determine coefficient a_n , we can apply the following derivation. Beginning with Definition 1.1, both sides of the Fourier series can be multiplied by $\cos(\omega m x)$, where $n, m \in \mathbb{Z}^+$.

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} (a_n \cos(\omega n x) + b_n \sin(\omega n x)) \\ f(x) \cos(\omega m x) &= \sum_{n=0}^{\infty} (a_n \cos(\omega n x) \cos(\omega m x) \\ &\quad + b_n \sin(\omega n x) \cos(\omega m x)) \end{aligned}$$

We can then integrate over period T :

$$\begin{aligned} & \int_0^T f(x) \cos(\omega m x) \, dx \\ &= \int_0^T \sum_{n=0}^{\infty} (a_n \cos(\omega n x) \cos(\omega m x) \\ &\quad + b_n \sin(\omega n x) \cos(\omega m x)) \, dx \end{aligned}$$

Because of distributive properties, and because a_n and b_n are constants:

$$\begin{aligned} & \int_0^T f(x) \cos(\omega m x) \, dx \\ &= \sum_{n=0}^{\infty} a_n \int_0^T \cos(\omega n x) \cos(\omega m x) \, dx \\ &\quad + \sum_{n=0}^{\infty} b_n \int_0^T \sin(\omega n x) \cos(\omega m x) \, dx \end{aligned}$$

Using Claim 2.1, we can isolate the cosine functions:

$$\begin{aligned} & \int_0^T f(x) \cos(\omega m x) \, dx \\ &= \sum_{n=0}^{\infty} a_n \int_0^T \cos(\omega n x) \cos(\omega m x) \, dx + 0 \end{aligned}$$

As proven in Claim 2.2, there are two possible cases for each term in the infinite summation.

When $m \neq n$:

$$\begin{aligned} & a_n \int_0^T \cos(\omega n x) \cos(\omega m x) \, dx \\ &= a_n(0) = 0 \end{aligned}$$

When $m = n$:

$$a_m \int_0^T \cos(\omega n x) \cos(\omega m x) \, dx = a_m \left(\frac{1}{2} T \right)$$

Evidently, the only the term that will contribute to the Fourier series is when $m = n$, because every other term would simply equate to 0. We can thus interchange m and n , as in the following:

$$\int_0^T f(x) \cos(\omega n x) dx = \frac{1}{2} a_n T$$

The equation for the subject a_n can be rearranged:

$$a_n = \frac{2}{T} \int_0^T f(x) \cos(\omega n x) dx \quad (8)$$

As demonstrated, the previous proofs have linked nicely with the derivation process, and a_n can now be solved for. It further shows how arbitrarily manipulating equations (*ie*: multiplying both sides by $\cos(\omega n x)$), despite seeming random, works perfectly. This arbitrary manipulation is used in many proofs for physics and mathematics, and it strikes me how powerful it can be.

Coefficient a_0

For the case $n = 0$, or a_0 :

$$\begin{aligned} \int_0^T f(x) dx &= \sum_{n=0}^{\infty} a_n \int_0^T \cos(\omega n x) dx \\ &= a_0 \int_0^T \cos(\omega(0)x) dx \\ &\quad + \sum_{n=1}^{\infty} a_n \int_0^T \cos(\omega n x) dx \end{aligned}$$

Using Claim 1:

$$\begin{aligned} \int_0^T f(x) dx &= a_0 \int_0^T \cos(\omega(0)x) dx + \sum_{n=1}^{\infty} a_n \int_0^T 0 dx \\ &= a_0 \int_0^T 1 dx \\ &= a_0 T \end{aligned}$$

And by rearranging the equation for the subject a_0 :

$$a_0 = \frac{1}{T} \int_0^T f(x) dx \quad (9)$$

What's interesting, is that this is just the average value of $f(x)$. And this makes sense, because graphically, a_0 is a vertical translation. As sine and cosine waves oscillate above and below a midline, which is the average value of the function (*eg*: in Figure 2, sine and cosine waves oscillate about the x -axis), this translation will make the Fourier series

oscillate about the average value of the modelled function.

Coefficient b_n

The derivation follows closely to the derivation for a_n . b_n is derived as:

$$b_n = \frac{2}{T} \int_0^T f(x) \sin(\omega n x) dx \quad (10)$$

Unlike the cosine function, there is no special case for $n = 0$, because $\sin 0 = 0$.

Mini Conclusion

We have derived all coefficient to the Fourier series, and it is summarized below.

$$a_0 = \frac{1}{T} \int_0^T f(x) dx$$

$$a_n = \frac{2}{T} \int_0^T f(x) \cos(\omega n x) dx$$

$$b_n = \frac{2}{T} \int_0^T f(x) \sin(\omega n x) dx$$

This can now be applied to approximate periodic functions. However, this equation made it clear that it is impossible to approximate unknown functions, as finding a_0 , a_n , and b_n is dependent on integrating the original function, as well as integrating the product of the function and a trigonometric function. Upon further study, I realized that approximating an unknown function would require algorithms and more complicated derivations with Fourier transform. Thus, a curve fitting software was used to find the polynomial first then Fourier's series is used to turn it into a periodic function. So, this investigation will show how the Fourier series can be used in conjunction with technology to model ECG readings.

3 Approximating ECG Readings

The modelled ECG reading will be denoted as a function $V(t)$ for voltage (mV), t for time (s). However, it became clear that using the Fourier series is not perfect, as many types of abnormal readings would entail irregular rhythms/periods that would not be able to be simulated with the series, which has consistent periods.

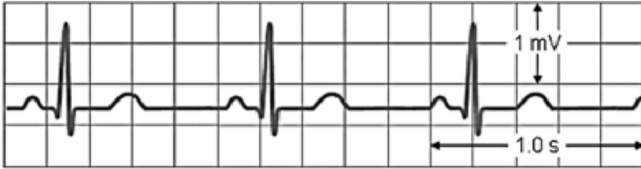


Fig. 8. ECG of healthy sinus rhythms. y -axis: voltage (mV), x -axis: time (s) (Kalra et al., 2018).

As healthy ECG datasets were not available online, and rather, only diagrams, a website (Rohangi, 2010) that converts images of graphs to plot points was utilized. The ECG reading in Figure 8 was hence converted to plot points, and the period was determined to be 0.951s. Excel was then chosen to fit the curve. As Excel polynomial trend lines only go to the 6th degree, the fitting was not perfect. It was thus approached like a piece-wise function and split it into curves and horizontal lines. In selecting the trend lines from Excel for my piece-wise function, I opted for the functions that have the greatest R^2 value (all >0.90), and ones that minimize the amount of functions required.

Piece-wise function:

$$V(t) = \begin{cases} 13926t^4 - 2208.5t^3 + 39.745t^2 + 4.1037t + 0.70318 & \text{if } 0 \leq t \leq 0.0879 \\ 0.703 & \text{if } 0.0879 \leq t \leq 0.160 \\ 762979t^4 - 584872t^3 + 166150t^2 - 20723t + 958.24 & \text{if } 0.160 < t \leq 0.237 \\ 0.703 & \text{if } 0.237 < t \leq 0.405 \\ -23.462t^2 + 23.091t - 4.8138 & \text{if } 0.405 < t \leq 0.572 \\ 0.703 & \text{if } 0.572 < t \leq 0.951 \end{cases} \quad (11)$$

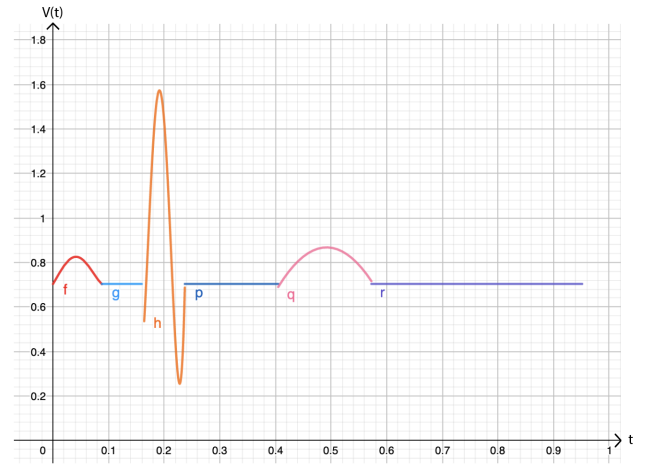


Fig. 9. Obtained piece-wise function graphed on GeoGebra (Hohenwarter et al., 2013), where $f, g, h, p, q,$ and r are representative of the sub-functions in Equation 11.

Evidently, the piece-wise function is not perfect and discontinuous at places (for instance, between g and h in Figure 9). However, it can be assumed that discontinuity would not heavily affect the produced Fourier series, as square waves (composed of discontinuous piece-wise functions, as in Figure 1) have been successfully, and smoothly, approximated with the Fourier series. Even so, if the curve fitting function is already somewhat inaccurate, as the peak of function h is smaller than that of the original ECG reading, the Fourier series will automatically be inaccurate as well. Using this method of curve fitting may not be highly reliable for such complex functions.

3.1 Coefficient a_0

$$a_0 = \frac{1}{T} \int_0^T V(t) dt$$

To account for the piece-wise function, the integral can be broken down to cater for each sub-function. It can be rewritten as:

$$a_0 = \frac{1}{0.951} \left(\int_0^{0.0879} V(t) dt + \int_{0.0879}^{0.160} V(t) dt + \int_{0.160}^{0.237} V(t) dt + \int_{0.237}^{0.405} V(t) dt + \int_{0.405}^{0.572} V(t) dt + \int_{0.572}^{0.951} V(t) dt \right)$$

Then, by substituting the respective sub-functions, we get:

$$a_0 = \frac{1}{0.951} \left(\int_0^{0.0879} (13926t^4 - 2208.5t^3 + 39.745t^2 + 4.1037t + 0.70318) dt + \int_{0.0879}^{0.160} 0.703 dt + \int_{0.160}^{0.237} (762979t^4 - 584872t^3 + 166150t^2 - 20723t + 958.24) dt + \int_{0.237}^{0.405} 0.703 dt + \int_{0.405}^{0.572} (-23.462t^2 + 23.091t - 4.8138) dt + \int_{0.572}^{0.951} 0.703 dt \right)$$

=0.748 mV (3 s.f.)

The approximate average value has now been obtained. It looks accurate, as it is slightly above the flat horizontal lines in the original piece-wise graph.

3.2 Coefficient a_n

Recall that:

$$a_n = \frac{2}{T} \int_0^T V(t) \cos(\omega nt) dt$$

Note that $\omega = \frac{2\pi}{T} = \frac{2\pi}{0.951}$. Substituting the respective piece-wise function we get:

$$a_n = \frac{2}{0.951} \left(\int_0^{0.0879} (13926t^4 - 2208.5t^3 + 39.745t^2 + 4.1037t + 0.70318) \cos(\omega nt) dt + \int_{0.0879}^{0.160} 0.703 \cos(\omega nt) dt + \int_{0.160}^{0.237} (762979t^4 - 584872t^3 + 166150t^2 - 20723t + 958.24) \cos(\omega nt) dt + \int_{0.237}^{0.405} 0.703 \cos(\omega nt) dt + \int_{0.405}^{0.572} (-23.462t^2 + 23.091t - 4.8138) \cos(\omega nt) dt + \int_{0.572}^{0.951} 0.703 \cos(\omega nt) dt \right)$$

Calculating such complex integrals is a very arduous, mechanical task, so it will simply be explained

with an example. For instance, function q from the piece-wise in Figure 9 will be calculated. It will then be expanded by distributing $\cos(\omega nt)$ and the integral:

$$\int_{0.405}^{0.572} (-23.462t^2 + 23.091t - 4.8138) \cos(\omega nt) dt$$

$$= - \int_{0.405}^{0.572} 23.462t^2 \cos(\omega nt) dt + \int_{0.405}^{0.572} 23.091t \cos(\omega nt) dt - \int_{0.405}^{0.572} 4.8138 \cos(\omega nt) dt$$

Integration by parts will be used, demonstrated with an example of one integral. First, the constant will be isolated, then integration by parts will be applied.

$$- \int_{0.405}^{0.572} 23.462t^2 \cos(\omega nt) dt$$

$$= -23.462 \left[t^2 \int \cos(\omega nt) dt - \int \frac{dt^2}{dt} \left(\int \cos(\omega nt) dt \right) dt \right]_{0.405}^{0.572}$$

$$= -23.462 \left[\left(t^2 \frac{1}{\omega n} \sin(\omega nt) \right) - \int 2t \left[\frac{1}{\omega n} \sin(\omega nt) \right] dt \right]_{0.405}^{0.572}$$

The last integral becomes another integration by parts problem:

$$- \int_{0.405}^{0.572} 2t \left(\frac{1}{\omega n} \sin(\omega nt) \right) dt$$

$$= - \frac{1}{\omega n} \left[t \int \sin(\omega nt) dt - \int \frac{dt}{dt} \left(\int \sin(\omega nt) dt \right) dt \right]_{0.405}^{0.572}$$

$$= - \frac{1}{\omega n} \left[t \left(- \frac{1}{\omega n} \cos(\omega nt) \right) - \int \left(- \frac{1}{\omega n} \cos(\omega nt) \right) dt \right]_{0.405}^{0.572}$$

$$= - \frac{1}{\omega n} \left[- \frac{t}{\omega n} \cos(\omega nt) + \frac{1}{\omega^2 n^2} \sin(\omega nt) \right]_{0.405}^{0.572}$$

$$= \left[\frac{t}{\omega^2 n^2} \cos(\omega nt) - \frac{1}{\omega^3 n^3} \sin(\omega nt) \right]_{0.405}^{0.572}$$

The calculations afterward is then just substitution of the limits and ω . This process can thus be repeated for each function and each integral. However, it is impractical to derive these equations by hand. Using technology such as Wolfram Alpha is much preferred, especially because of the degrees of the polynomials and how many sub-functions of the piecewise function are involved.

a_n is thus calculated as the following:

$$\begin{aligned}
a_n = & \frac{2}{0.951} \left(\frac{1}{n^5} \left[-0.0940106n^3 + (-0.0519801n^3 - 8.46377n) \cos(0.580749n) \right. \right. \\
& + (0.106316n^4 - 0.713945n^2 + 26.5486) \sin(0.580749n) - 6.95427n \left. \right] \\
& + \frac{1}{n} \left[0.106404 \sin(1.05711n) - 0.106404 \sin(0.580749n) \right] \\
& + \frac{1}{n^5} \left[-0.0590601n^4 \sin(1.05711n) + 0.103965n^4 \sin(1.56584n) \right. \\
& + (-0.629494n^3 - 304.071n) \cos(1.05711n) + (2.39033n^3 - 435.904n) \cos(1.56584n) \\
& + 18.0652n^2 \sin(1.05711n) - 51.5989n^2 \sin(1.56584n) - 1454.54 \sin(1.05711n) \\
& + 1454.54 \sin(1.56584n) \left. \right] + \frac{1}{n} \left[0.106404 \sin(2.6758n) - 0.106404 \sin(1.56584n) \right] \\
& + \frac{1}{n^3} \left[-0.104391n^2 \sin(2.6758n) + 0.108653n^2 \sin(3.77916n) \right. \\
& - 0.162704 \sin(2.6758n) + 0.162704 \sin(3.77916n) \\
& - 0.093623n \cos(2.6758n) - 0.085897n \cos(3.77916n) \left. \right] \\
& + \frac{1}{n} \left[0.106404 \sin(6.28319n) - 0.106404 \sin(3.77916n) \right] \left. \right)
\end{aligned}$$

3.3 Coefficient b_n

$$b_n = \frac{2}{T} \int_0^T V(t) \sin \omega n t \, dt$$

Following the same process as for coefficient a_n , we get the following equation for b_n :

$$\begin{aligned}
b_n = & \frac{2}{0.951} \left(\frac{1}{n^5} \left[0.106431n^4 + (-0.0519801n^3 - 8.46377n) \sin(0.580749n) \right. \right. \\
& - 0.275622n^2 + (-0.106316n^4 + 0.713945n^2 - 26.5486) \cos(0.580749n) + 26.5486 \left. \right] \\
& + \frac{1}{n} \left[0.106404 \cos(0.580749n) - 0.106404 \cos(1.05711n) \right] \\
& + \frac{1}{n^5} \left[-0.629494n^3 \sin(1.05711n) + 2.39033n^3 \sin(1.56584n) \right. \\
& + (0.0590601n^4 - 18.0652n^2 + 1454.54) \cos(1.05711n) \\
& + (-0.103965n^4 + 51.5989n^2 - 1454.54) \cos(1.56584n) - 304.071n \sin(1.05711n) \\
& - 435.904n \sin(1.56584n) \left. \right] + \frac{1}{n} \left[0.106404 \cos(1.56584n) - 0.106404 \cos(2.6758n) \right] \\
& + \frac{1}{n^3} \left[(0.104391n^2 + 0.162704) \cos(2.6758n) + (-0.108653n^2 - 0.162704) \cos(3.77916n) \right. \\
& - 0.093623n \sin(2.6758n) - 0.085897n \sin(3.77916n) \left. \right] \\
& + \frac{1}{n} \left[0.106404 \cos(3.77916n) - 0.106404 \cos(6.28319n) \right] \left. \right)
\end{aligned}$$

3.4 Collating Coefficient

Now that we have the general equations for each coefficient we can relate back to Equation 1:

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(\omega n x) + b_n \sin(\omega n x))$$

We can substitute the equations for each coefficient back to the overall Fourier series to find the expansion. The substitution process will not be shown, as evidently, they are very lengthy equations. I programmed a mini calculator that solves these coefficients to ease the process further, highlighting the need for using technology in conjunction to these calculations.

A table has been created with each coefficient rounded to 3 significant figures. This degree of accuracy was chosen because I acknowledge that the data points used to curve fit the piece-wise function is not perfectly accurate in the first place, so the resulting coefficients would as a result not be very accurate anyway. So, it is rather pointless to approximate to more significant figures. Especially because coefficients act as a vertical stretch, considering the scale of the coefficients decimal places typically beyond 3 significant figure would be so small that it is negligible in relation to the scale of the y -axis that is being worked with. However, rounding to any less than 3 significant figure may oversimplify results and lead to too much error. Further, it makes it more manageable to input into graphing calculators, especially when working with the summation of numerous trigonometric functions. Only 20 coefficients were calculated due to limited processing power for graphing.

n	a _n	b _n
0	0.748	0
1	-0.0136	0.0217
2	0.0161	0.041
3	-0.0558	-0.0186
4	0.0318	-0.0226
5	0.0333	-0.00594
6	0.0298	0.0587
7	-0.0419	0.0413
8	-0.0504	-0.0177
9	0.00689	-0.0457
10	0.0463	-0.00369
11	0.0131	0.0542
12	-0.0482	0.0258
13	-0.0402	-0.034
14	0.0207	-0.0443
15	0.0436	0.00824
16	0.000526	0.0427
17	-0.0361	0.00879
18	-0.0171	-0.0275
19	0.0201	-0.0195
20	0.0189	0.0127

Table 1. First 20 n values for a_n and b_n coefficients

```
import math
n=0
def eq1(n):
    var1 = -0.0940186*n**3 + (-0.0519881*n**3 - 8.46377*n)*math.cos(0.580749*n)
    var2 = (0.186316*n**4 - 0.713945*n**2 + 26.5486)*math.sin(0.580749*n) - 6.95427*n
    return (var1+var2)/n**5
def eq2(n):
    return (0.186404*math.sin(1.05711*n) - 0.186404*math.sin(0.580749*n))/n
def eq3(n):
    var1 = -0.0598081*n**4*math.sin(1.05711*n) + 0.183965*n**4*math.sin(1.56584*n)
    var2 = (-0.629494*n**3 - 384.071*n)*math.cos(1.05711*n) + (2.39833*n**3 - 435.904*n)*math.cos(1.56584*n)
    var3 = 18.0652*n**2*math.sin(1.05711*n) - 51.5989*n**2*math.sin(1.56584*n) - 1454.54*math.sin(1.05711*n)
    var4 = 1454.54*math.sin(1.56584*n)
    return (var1+var2+var3+var4)/n**5
def eq4(n):
    return (0.186404*math.sin(2.6758*n) - 0.186404*math.sin(1.56584*n))/n
def eq5(n):
    var1 = -0.184391*n**2*math.sin(2.6758*n) + 0.186653*n**2*math.sin(3.77916*n) - 0.162704*math.sin(2.6758*n)
    var2 = 0.162704*math.sin(3.77916*n) - 0.093623*n*math.cos(2.6758*n) - 0.085897*n*math.cos(3.77916*n)
    return (var1+var2)/n**3
def eq6(n):
    return (0.186404*math.sin(6.28319*n) - 0.186404*math.sin(3.77916*n))/n
def a_n(n):
    return (2/0.951)*(eq1(n)+eq2(n)+eq3(n)+eq4(n)+eq5(n)+eq6(n))
for i in range(1, 21):
    print(a_n(i))
```

Fig. 10. Mini calculator programmed in Python for coefficient a_n (left), and the calculated values (right).

3.5 Final Results and Evaluation

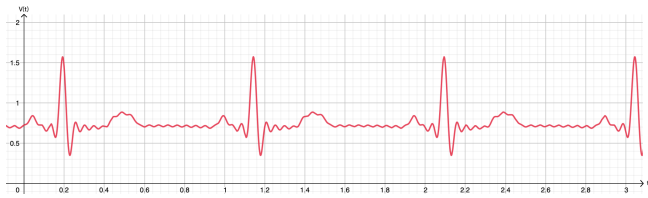


Fig. 11. The Fourier series of the healthy ECG reading with the first 20 n values graphed on GeoGebra (Hohenwarter et al., 2013).

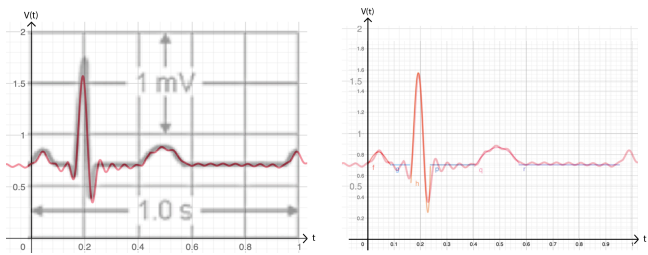


Fig. 12. Overlay of the Fourier series in Figure 11 on the original ECG reading in Figure 8 (Kalra et al., 2018) (left) and on the piecewise function in Figure 9 (right).

Using the labels in Figure 6 to refer to parts of the ECG reading, the overall shape closely resembles the original graph. The P, Q, R, S and T waves are each situated around the right coordinates, and the period is correct. But, it is also clear that it is not perfect, particularly pertaining the R wave, where the R maximum has a shorter amplitude than the original ECG. Perhaps because of the nature of sine and cosine functions and working with only 20 coefficients the supposed flat horizontal lines are still visibly curved, and the P and T waves are composed of many peaks. If such function is used for analysis, it would lack the precision required and may actually appear like an abnormal and unhealthy ECG reading.

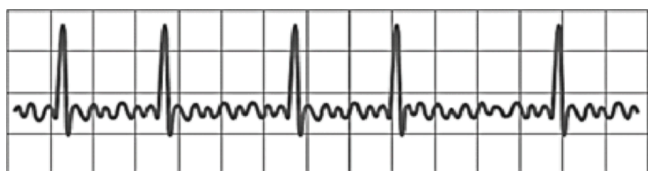


Fig. 13. ECG reading of atrial fibrillation a heart rhythm disorder (Kalra et al., 2018).

It likely requires a greater amount of coefficient to approximate it more closely to the original piecewise function in Figure 9, as the original definition of the Fourier series sets n to infinity. With this resemblance, however, in the future, the Fourier series for healthy ECGs and unhealthy ECGs can

be used to conduct automatic preliminary analysis without excessive monitoring.

While the results were certainly interesting and somewhat impressive, using the Fourier series to approximate complex functions, particularly one that requires precision, is not ideal. While determining coefficient is simply a substitution process (given that the equation is known), the integration can get tricky for higher order polynomials and piecewise functions. Further, the Fourier series is limited in approximating unknown functions, and using curve fitting tools then applying the series is inefficient and opens up more room for error.

Considering that one of the most significant flaws are the heights of the peak, which is only attributed to imperfect curve fitting rather than the Fourier series, it demonstrates that perhaps it could have been significantly bettered if higher order polynomials and more precise data points were used (contrary to selecting pixels on a graph as data points).

A potential solution, Fourier transform, is the opposite of the Fourier series, where functions are decomposed into its constituent trigonometric functions. It is the more common and efficient basis for many algorithms such as Fast Fourier Transform to quickly approximate unknown functions. It is used in the technologies for electricity, waves, etc., being an essential mathematical tool that revolutionized many aspects of science. In future endeavors, I should expand on Fourier series to understand Fourier transform and its related algorithms, and hopefully yield results more efficiently.

In the future, I can also use the Fourier series, or Fourier transform, for ECGs of different heart conditions, and determine a numerical based analysis and comparison between the period, a_0 , a_n , and b_n coefficient of healthy and unhealthy/abnormal ECGs, which may be used in an algorithm that can automatically determine certain conditions. However, this is only given that the Fourier series or transform is indeed precise enough to do so. This can be ensured by using statistical distance analyses of how similar the calculated Fourier series is to the original ECG reading. Once this can be ensured, this mathematical investigation can be fur-

thered after modelling with big data from medical centers, and perhaps by applying a machine learning algorithm to help doctors identify less obvious heart conditions, or to identify signs that a heart condition will be imminent. This algorithm can be eventually stored in the ECG and doctors could be immediately notified reducing room for human errors and will not require consistent monitoring to detect a heart problem. This method hence has a great potential in medical applications.

References

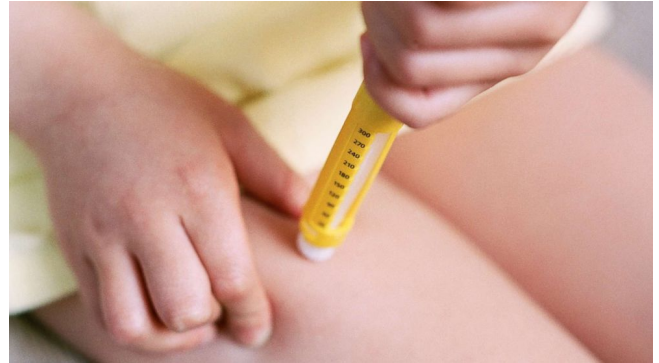
- 3Blue1Brown. (2019, June). *But what is a Fourier series? From heat flow to circle drawings | DE4*. https://www.youtube.com/watch?v=r6sGWTCMz2k&t=113s&ab_channel=3Blue1Brown.
- Anderson, J., & DiCarlo, S. E. (2000). "Virtual" experiment for understanding the electrocardiogram and the mean electrical axis. *Advances in physiology education*, 23(1), S1–17.
- Atkielski, A. (2007, January). *Schematic diagram of normal sinus rhythm for a human heart as seen on ECG (with English labels)*. <https://commons.wikimedia.org/wiki/File:SinusRhythmLabels.svg>.
- Cheever, E. (2019). *Derivation of Fourier Series*. <https://lpsa.swarthmore.edu/Fourier/Series/DerFS.html>. Swarthmore.
- Chou, K.-S. (n.d.). *Chapter 1 Fourier Series*. https://www.math.cuhk.edu.hk/course_builder/1415/math3060/Chapter\%201.\%20Fourier\%20series\%20.pdf.
- Hohenwarter, M., Borchers, M., Ancsin, G., Bencze, B., Blossier, M., Delobelle, A., ... Sturr, G. (2013, dec). *GeoGebra 4.4*. (<http://www.geogebra.org>)
- Kalra, A., Lowe, A., & Al-Jumaily, A. (2018). Critical review of electrocardiography measurement systems and technology. *Measurement Science and Technology*, 30(1), 012001.
- Khan, S. (2008, January). *Signals and systems*. <https://www.khanacademy.org/science/electrical-engineering/ee-signals#ee-fourier-series>.
- Price, D. (2010). *How to read an Electrocardiogram (ECG). Part One: Basic principles of the ECG. The normal ECG*. Retrieved from <http://www.southsudanmedicaljournal.com/archive/may-2010/how-to-read-an-electrocardiogram-ecg.-part-one-basic-principles-of-the-ecg.-the-normal-ecg.html>
- Rohangi, A. (2010). *WebPlotDigitizer - Copyright 2010-2019 Ankit Rohatgi*. <https://apps.automeris.io/wpd/>.

Design Technology: EpiPen

Stanley Ip 葉宇軒

1 Design Opportunity

An estimated 1 in 13 children (McCray, 2019) experience life-threatening anaphylaxis (Turner). An epinephrine auto-injector (EAI) is a consumer device that can relieve these reactions. Though, recent EAIs have flaws that can cause injury in children and inaccessibility to low-income families.



1.1 Injury to Children

EAIs have caused up to 6,806 injuries through cuts and lacerations in children aged 3-10, many who self-administered (in America) (Anshien). Recent EAIs are not suitable for children who self-administer because of:

Poor Use Guidance

Recent EAIs instruct the user to 'swing' the device during use. This action increases movement during injection, increasing chances of injury. Through demonstration with a practise pen with a trained nurse, it was found that only a firm push was necessary.

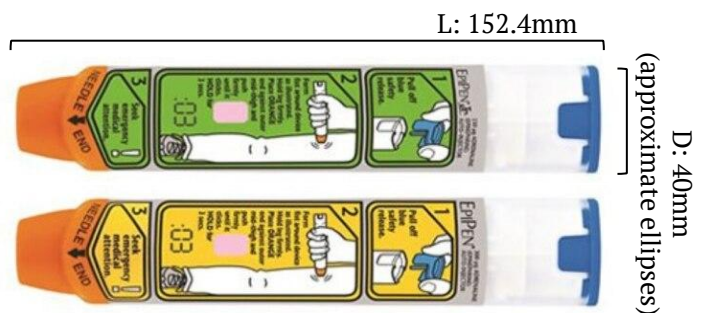
This action is suggested on the labels of some EAIs

Small and difficult to read

The above however may be a more safe label

Ergonomics and Anthropometrics

EAIs such as the EpiPen are the same for both adult and children. Considering the physiological factors for children in comparison to adults:



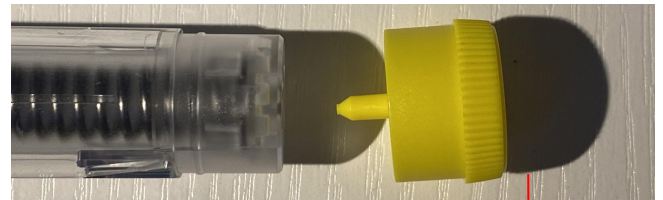
The model was made to approximate the dimensions of an EAI). As seen, the fingers of the 8 year old user are not clapsed, preventing a strong fist-grip.

Factor	Adult	Child (~10 years)
Average Grip Strength	48-51kg (male) 26-29kg (female)	12.9-18.3kg (male & female)
Average Grip Size (L x W)	19.3 x 8.9cm (male) 17.2 x 7.8cm (female)	15.7 x 6.5cm (male & female)

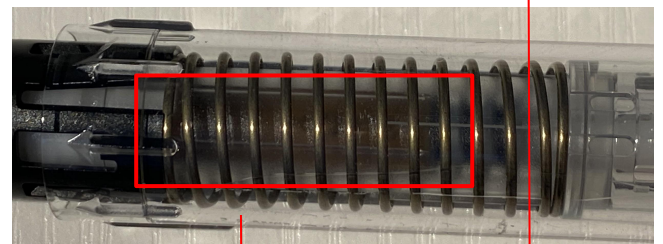
1.2 Accessibility

At up to 300 USD, an EAI can cost half of a month's rent for a low income family in the US (Rudden). In the long term, these costs can add up due to perishability, and the one use nature of EAIs.

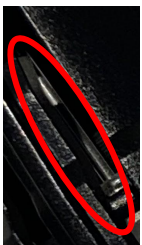
The prices of EAIs are marked up by up to 4000%, causing inaccessibility to low income families. The full cost of production is \$6-8 (label, other internals) (Mercury News) – and around \$2, or 28-33% of full cost is used to produce syringe and solution.



Syringe (25¢), Solution (\$1)



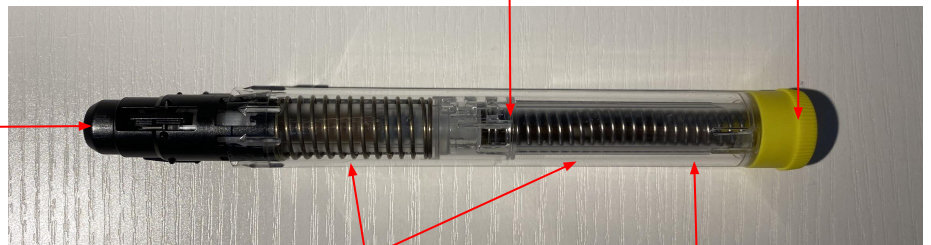
Lock (7¢)



Needle (25¢)



Cover (45¢)



Spring(s) (10¢)

Shell (20¢)

1.3 Product Analysis



EpiPen



Auvi-Q

Advantages

- Elliptical, ergonomic shape
- Compact, easy to store/carry
- Economic materials (mainly plastic construction)
- Economic manufacturing (likely to be injection moulded)

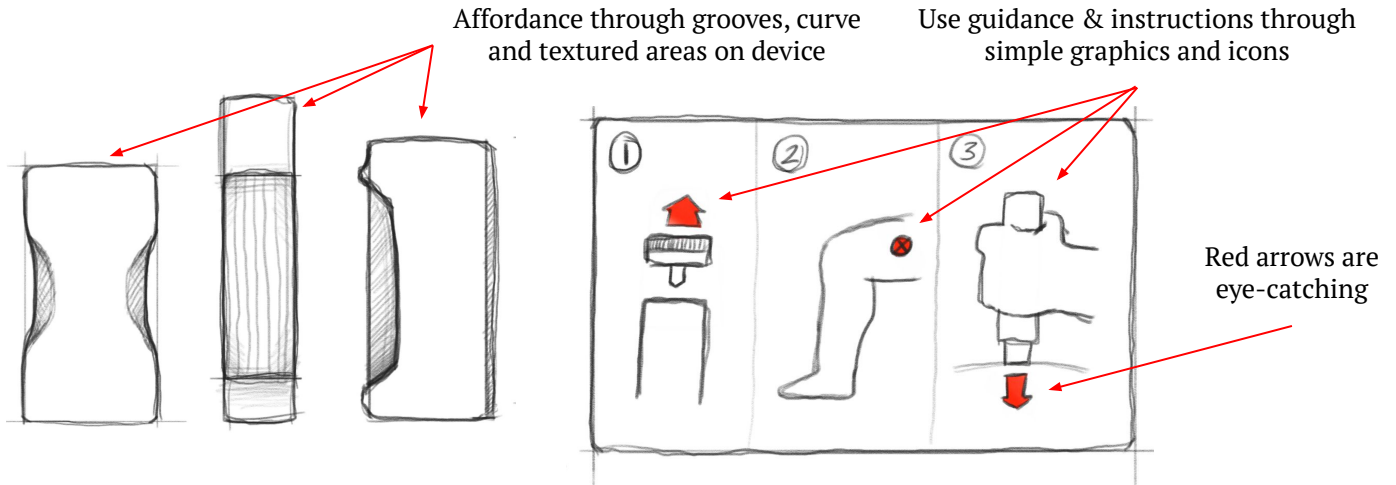
- Compact, easy to store/carry
- Automatic injection/retraction of needle (through electronics)
- Economic manufacturing (case likely to be injection moulded)

Disadvantages

- Single use and highly expensive to purchase (~300 USD)
- Use guidance is cluttered and little achieved
- Lack of adjustability/range of sizes

- Rectangular shape is not ergonomic to hold
- Single use + implementation of electronics increase cost to consumers (~2500 USD)
- Lack of adjustability/range of sizes

1.3 Morphological Synthesis



2 Brief

EAs are hazardous for child use (Anshien). Namely, poor use guidance and ergonomics can increase injury, complicating already difficult health situations. EAs are expensive. Their hefty price tags makes it a premium for low-income families (Rudden), for the ability to save a life during emergency allergic reactions.

2.1 Persona Mapping

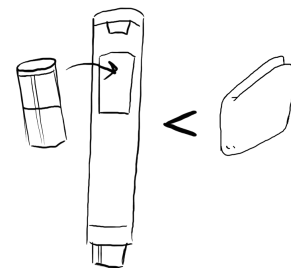
Age:	7
Title:	Elementary school student
Household Income:	~45,000 (annually) (ASPE)
Allergic to:	Bee stings
Sensitivity:	Type I (immediate hypersensitivity – anaphylactic reaction)
Details:	Must replace EAs yearly, and keep multiple at home and on carry. Has prior understanding to self-administering EAs.

2.2 Goal



A. **Able to be self-administered** for children aged 6 and up. At age 6, children are expected to have developed self-care skills & knowledge, adequate hand dexterity and hand-eye coordination (Buffalo Hearing). Ages below 6 will require adult guidance.

- Improved ergonomics & anthropometrics?
- Improved use guidance?
- Improved affordance?




B. **More affordable** to families from varying financial backgrounds.

- Reusability?
- Part reduction?
- Implementation of trade-ins?

2.3 Constraints

- Limited access to substances (solution & epinephrine)
- Limited access to target audience for testing
 - This will limit opportunities for participatory design and hence affect the quality and effectiveness of the product
- Limited access to means of manufacturing and economies of scale.
 - This will significantly increase the costs of production, and hence the price of the product
 - Limited economies of scale also limits the number of units that can be produced

3 Design Specification

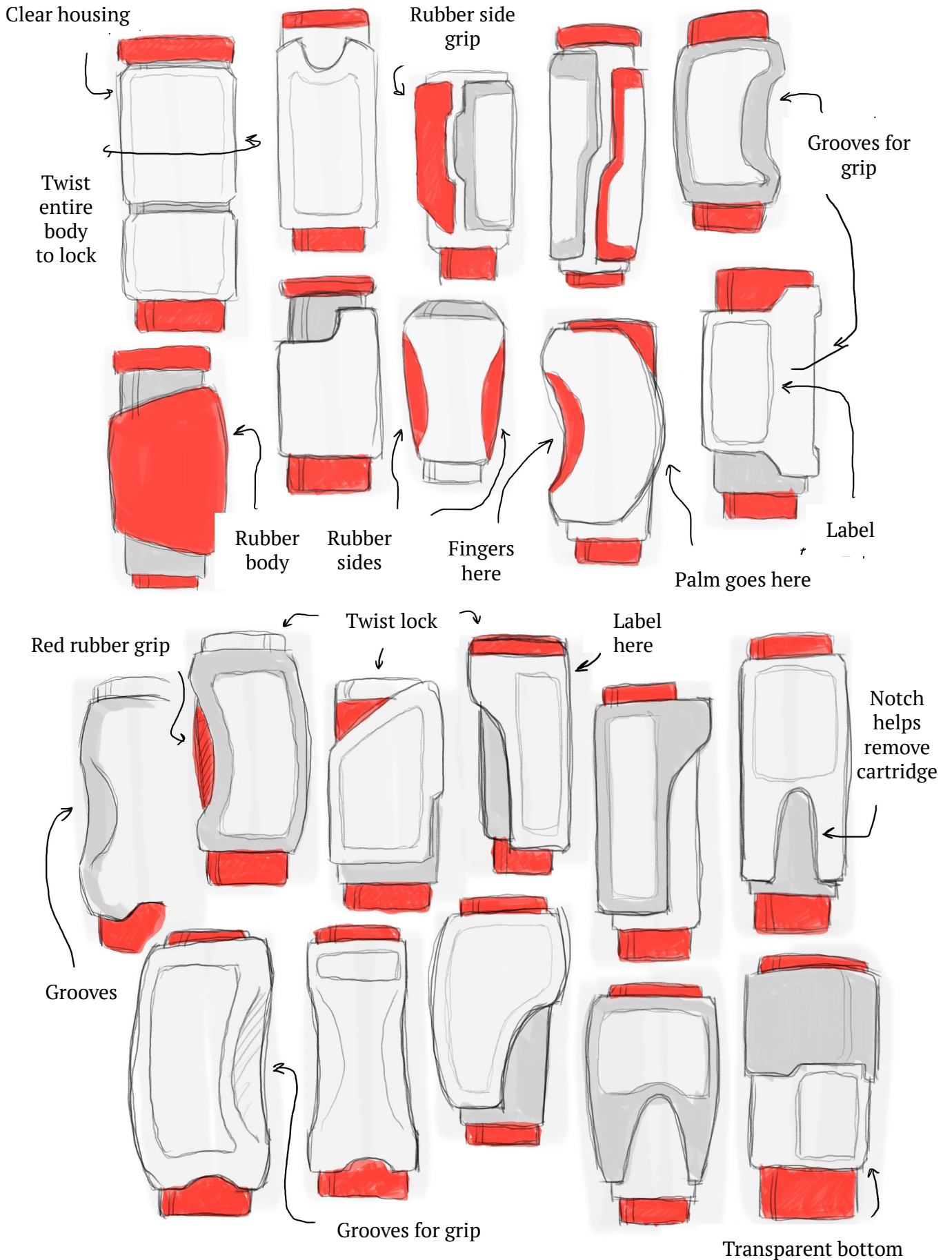
Specifications	Priority	Description	Justification
Aesthetics	5	<p>5.1: The device must have bold color indicators such as red & orange.</p>  <p>5.2: Graphics (instructions) must be simple and easy to understand (language and graphics), and utilize color-coding.</p>	<p>5.1: The red and orange stimulate danger and alertness (Pancare) in important areas and can work in conjunction with graphics.</p> <p>5.2: Poor use guidance leads to increased injury in children. To improve, simplicity allows the graphics to be understood quicker and by a larger age range, and color-coding will allow important parts of the device to be identified.</p>
Target Audience	6	<p>6.1: Children aged 6-12, who have less financially stable backgrounds and suffer from Type 1 anaphylactic shock.</p>	<p>6.1: Children are at a higher risk of injury when self-administering EAIs because of poor use guidance and ergonomics. The age range was selected to target users who would more likely use an EAI with little to no adult guidance.</p> <p>Low income families are not able to access EAIs due to their high prices, which can be life-threatening.</p>
Function	2	<p>2.1: 0.3mg of solution will be delivered per use of the EAI.</p> <p>2.2: The mechanism must be fast-acting and be delivered intentionally.</p> <p>2.3: Doses of solution must be replaceable - while the housing and mechanism can be reused.</p>	<p>2.1: 0.3mg is the safe dosage for children aged 6-12</p> <p>2.2: Anaphylaxis can escalate quickly (Harvard Health). A fast-acting mechanism will prevent further escalation and harm to the patient.</p> <p>2.3: Replaceable doses will allow the firing mechanism to be reused, which makes up 67~72% the price of some EAIs (Mercury News). This will reduce the long-term cost to consumers and increase accessibility.</p>

Specifications	Priority	Description	Justification
Ergonomics	3	<p>3.1: The housing must be mostly rounded/elliptical, with a diameter of ~30-35mm</p> <p>3.2: The housing must include grooves/textured areas and must be ambidextrous.</p>	<p>3.1: A diameter of ~30-35mm is found to be most comfortable for children aged 6-12 (Brown). This will also allow for the highest exertion of strength from the user, reducing chances of injury.</p> <p>3.2: Grooves and textured areas increases comfort and grip of the device, and suggests how the device should be held for safety (affordance).</p>
Safety	1	<p>1.1: The device must have a safety lock that can prevent misfires when not in use. It must be identifiable and easy to operate.</p> <p>1.2: A cover/cap must be provided per cartridge that can protect its internals (needle & syringe) before and after use.</p>	<p>1.1: EAI misfires can cause severe bodily damage (Tirell) and interrupt use during emergency situations. As anaphylaxis escalates quickly, the lock must be accessible to prevent further escalation and harm.</p> <p>1.2: As the needle will puncture beneath skin during use, it must remain sterile and clean to prevent further health complications. Injuries due to exposed needles after use can cause bloodborne infections diseases (CDC).</p>
Size	7	<p>7.1: Height: 150mm Diameter: ~30-35mm</p>	<p>7.1: The housing must accommodate mechanisms whilst remaining portable and easy to store, so that it may be kept on-hand for emergencies. The diameter addresses hand ergonomics.</p>
Materials	8	<p>8.1: The housing and cartridge will be made of HIPs plastic</p> <p>8.2: Metal components will be made of steel.</p>	<p>8.1: HIPs is lightweight and impact resistant (Blackwell Plastics). It also allows for methods such as injection moulding and is cheap and easy to acquire.</p> <p>8.2: Steel is hard, tough and resistant to corrosion, allowing the product's mechanisms to remain sterile and function for longer (Steel Metal Properties).</p>
Price	4	<p>4.1: Price for device will be in the range of 200-250 USD. Each cartridge will be priced in the range of 60-70 USD.</p>	<p>4.1: Current EAI's can cost up to 300 USD per device (Rudden). The proposed price range is more affordable, especially in the long-term as cartridges are replaceable (cartridges are priced at 28~33% of total price).</p> <p>The price is also needed to justify the purchase of materials, and the lack of economies-of-scale in manufacturing.</p>

Specifications	Priority	Description	Justification
Manufacturing	9	<p>9.1: Injection moulding will be used to manufacture the housing.</p> <p>9.2: Cartridges will be produced with batch production.</p>	<p>9.1: Since the housing is made largely from thermoplastics (e.g. HIPs) and has a hollow form, injection moulding is suitable.</p> <p>9.2: Batch production will be used for the cartridges, as the demand for cartridges applies to a niche demographic and epinephrine is perishable.</p> <p>9.1, 9.2: The methods chosen allow for economic manufacturing, lowering cost to consumers.</p>
Market Analysis	12	12.1: The EAI market is monopolized by distinct brands, with little competition and stimulation for innovation.	12.1: There are just more than 6 manufacturers worldwide. One manufacturer holds up to 72% market share - which has subsequently caused the 480% markup in prices of EAIs (Herman)
Environmental Considerations	11	11.1: The cartridge will be disposed after one use. It will not be recycled or disposed of conventionally.	11.1: The cartridge will be in direct contact with organic material and chemicals and can injure others during disposal. Epinephrine is categorized as a hazardous waste (EPA) and is a contaminant.
Constraints	10	<p>10.1: Very limited access to epinephrine</p> <p>Limited access to means of manufacturing and capital (injection moulding)</p> <p>Only a working prototype may be produced (note that due to this restriction, only PLA plastic could be used via FDM manufacturing)</p>	<p>10.1: Epinephrine may only be obtained from prescription (Hong Kong Journal of Paediatrics)</p> <p>Capital such as injection moulders are not available in school facilities.</p>

4 Initial Designs

4.1 Device



4.2 Label & Graphics

'E, E' → EMPHASIZE 'E'!

EPINEPHRINE

EPINEPHRINE

EPINEPHRINE

epinephrine

EPINEPHRINE

EPINEPHRINE

epinephrine

← SIMPLE

EPINEPHRINE

EPINEPHRINE

EPINEPHRINE

EPINEPHRINE

EPINEPHRINE

EPINEPHRINE

EPINEPHRINE

SHORTER STEMS

SIMPLE, SHARP EDGES

BLOCKY

ROUNDED EDGES.

BOLD, PLAYFUL, FONT

SIMPLE, SHARP, WIDER.

(fonts are typed)

MAJOR VERBS IN BLUE

1. TWIST CAP

MORE CLEAR

2. PUT AGAINST LEG, SIDE OF LEG.

PUSH FIRMLY FOR

3 SECONDS.

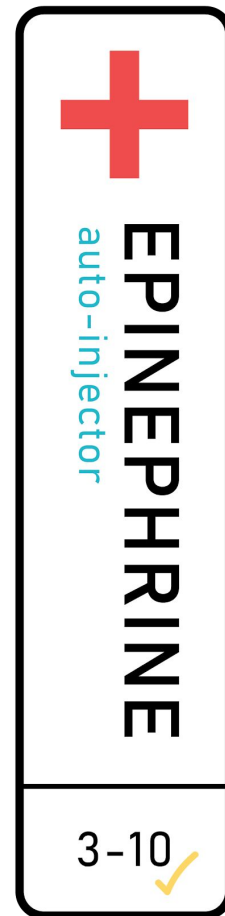
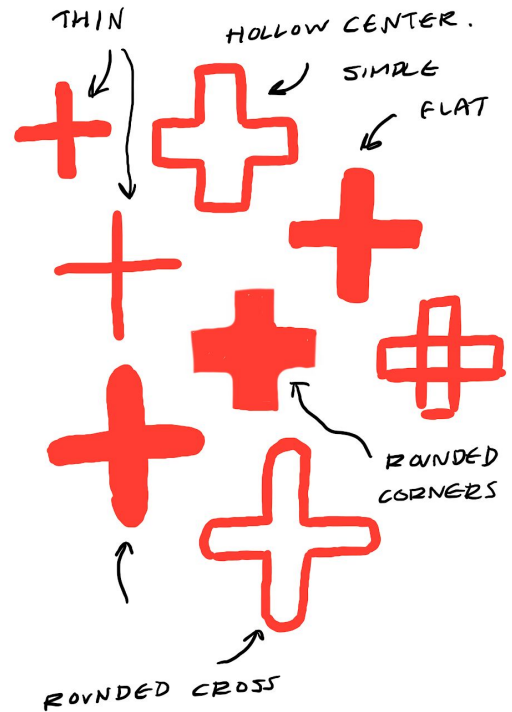
KEYWORDS IN YELLOW

3. REMOVE

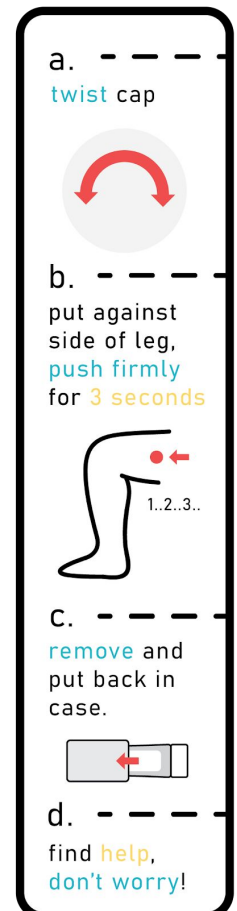
MIGHT BE CONFUSION BETWEEN NUMBERS?

4. SEEK HELP

USE SIMPLER VOCABULARY 'FIND'



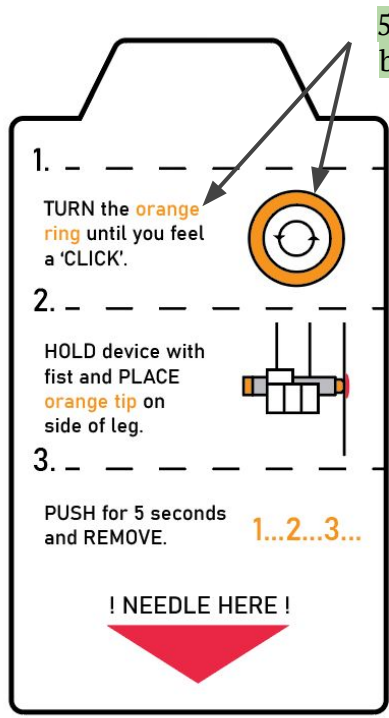
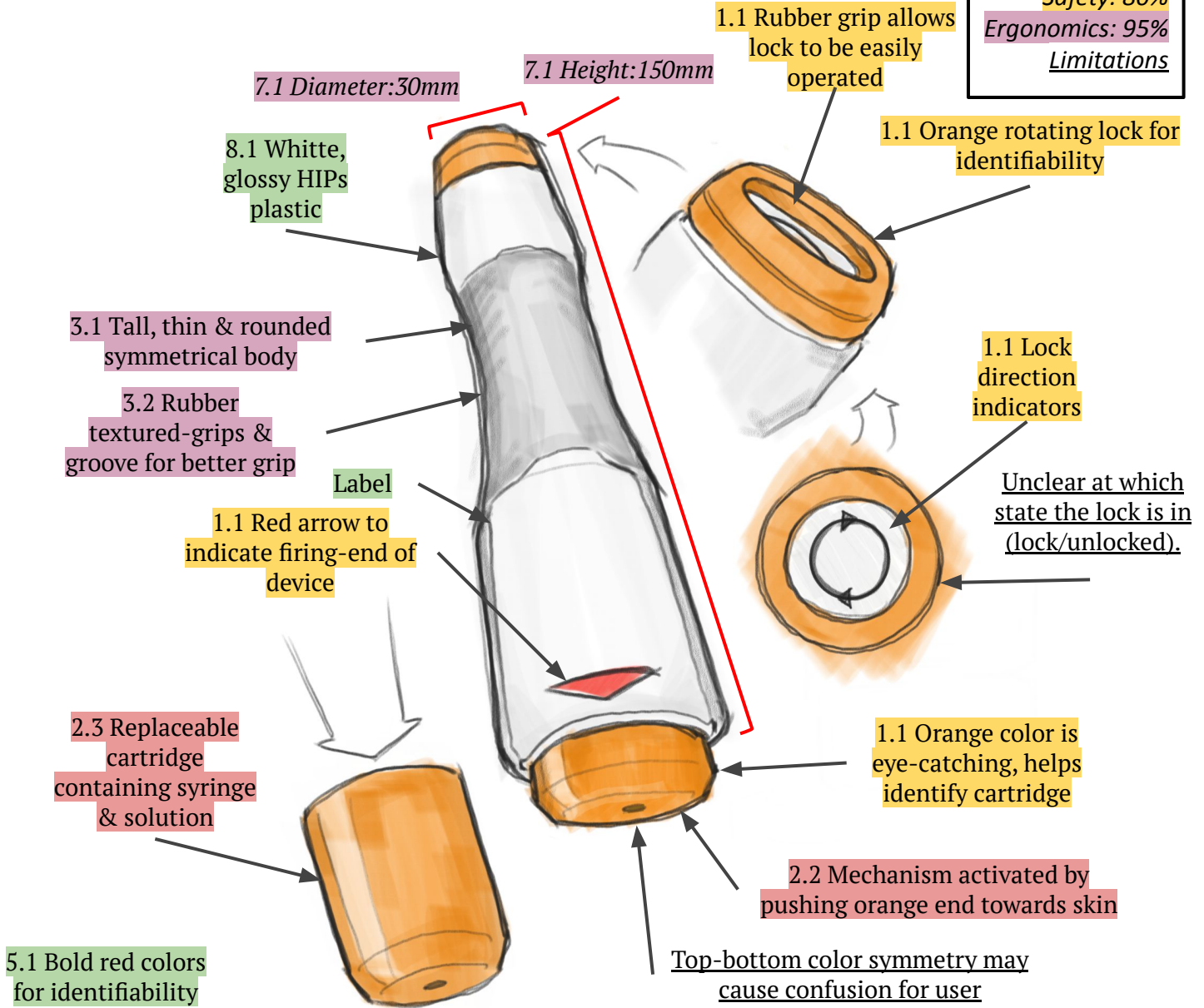
Front label



Instructions label

4.3 Design #1

Function: 90%
 Aesthetics: 90%
 Safety: 80%
 Ergonomics: 95%
Limitations



5.2 Color-coding for better use guidance

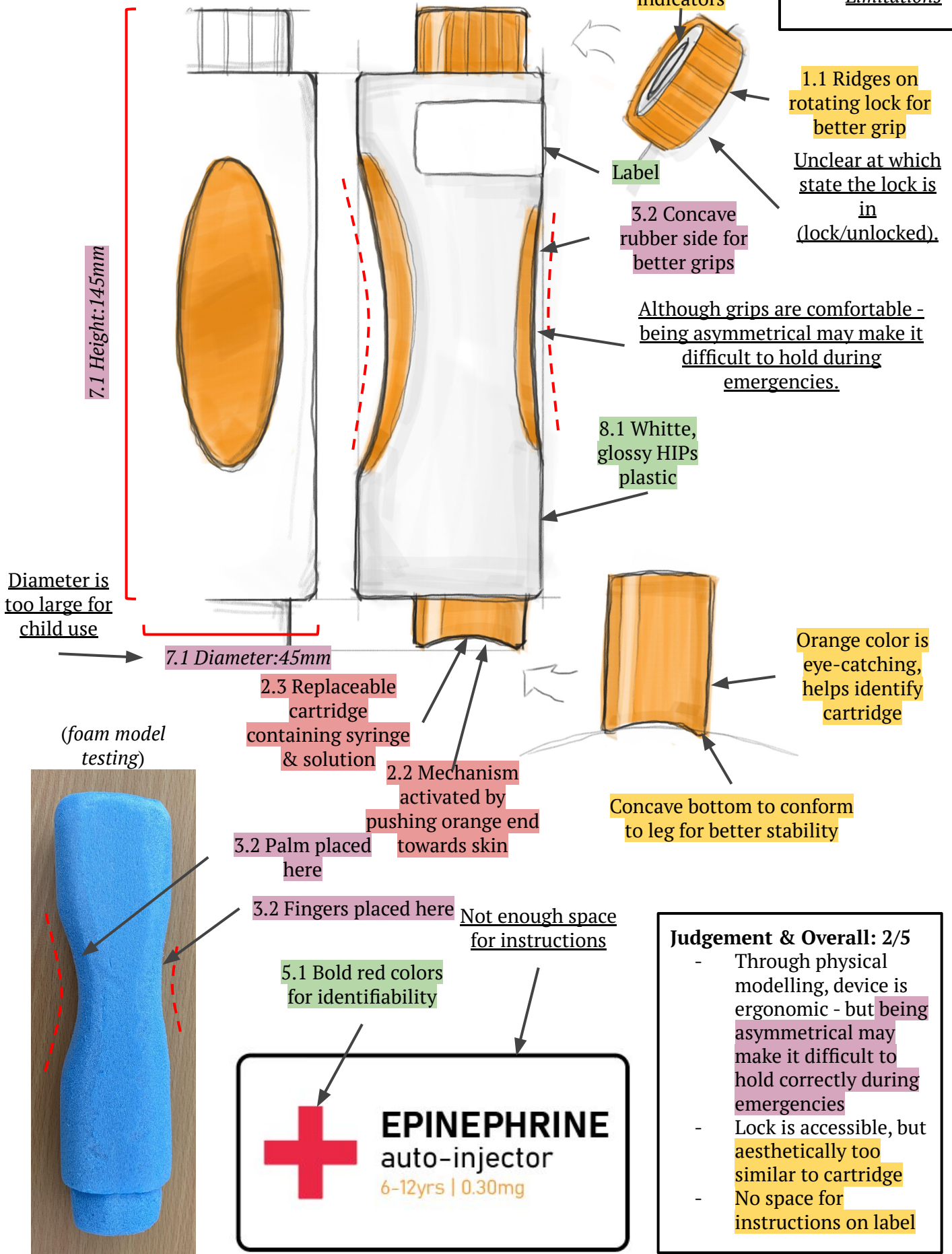
5.1 Simple, minimal graphics for better understandability

Judgement & Overall: 4/5

- Device is ergonomic and easy to handle
- Grooves on side assist user in holding the device in a stronger position
- Lock mechanism is identifiable, but too similar aesthetically to cartridge
- Label may be too small to read

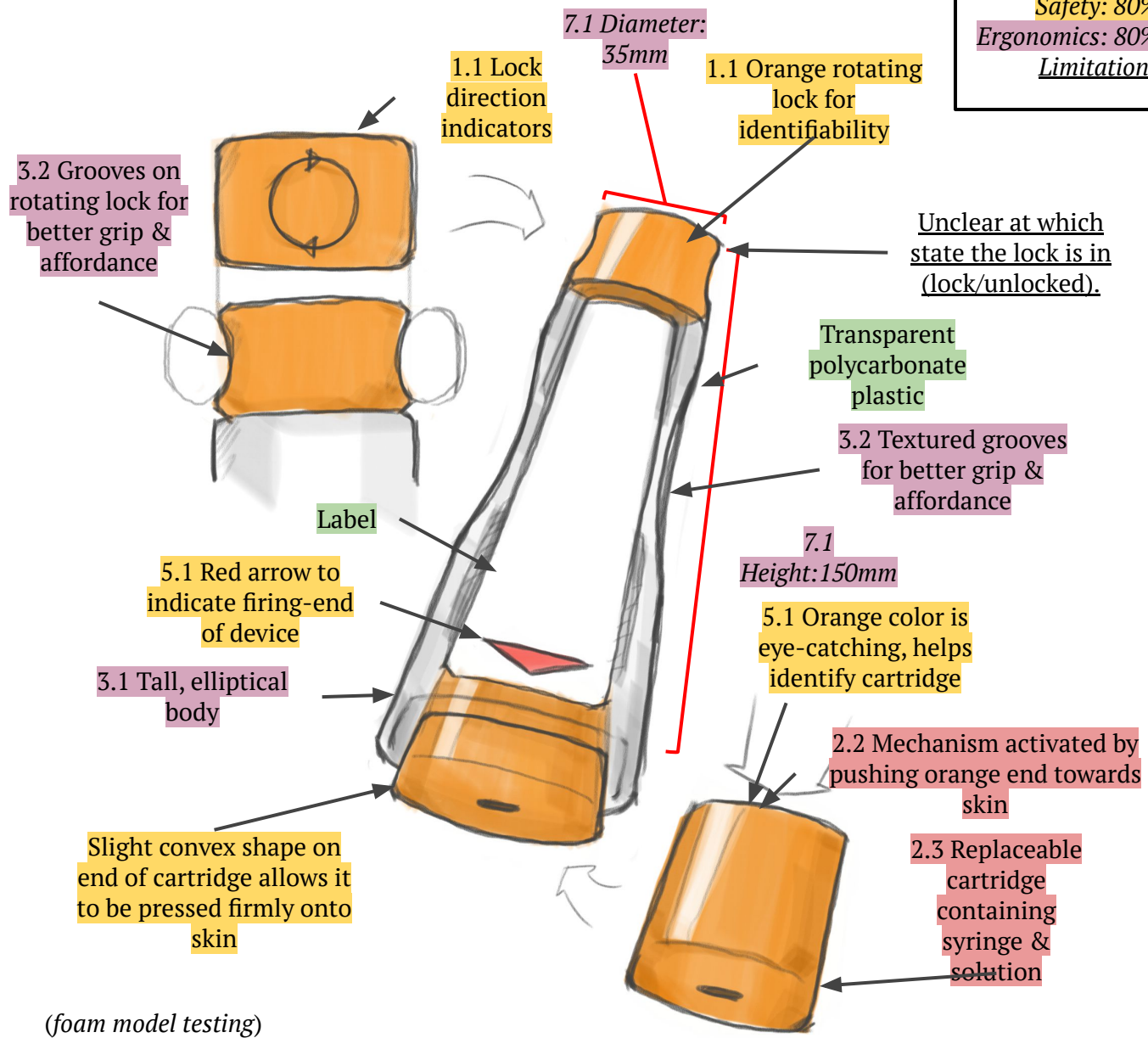
4.4 Design #2

Function: 90%
 Aesthetics: 80%
 Safety: 50%
 Ergonomics: 70%
 Limitations

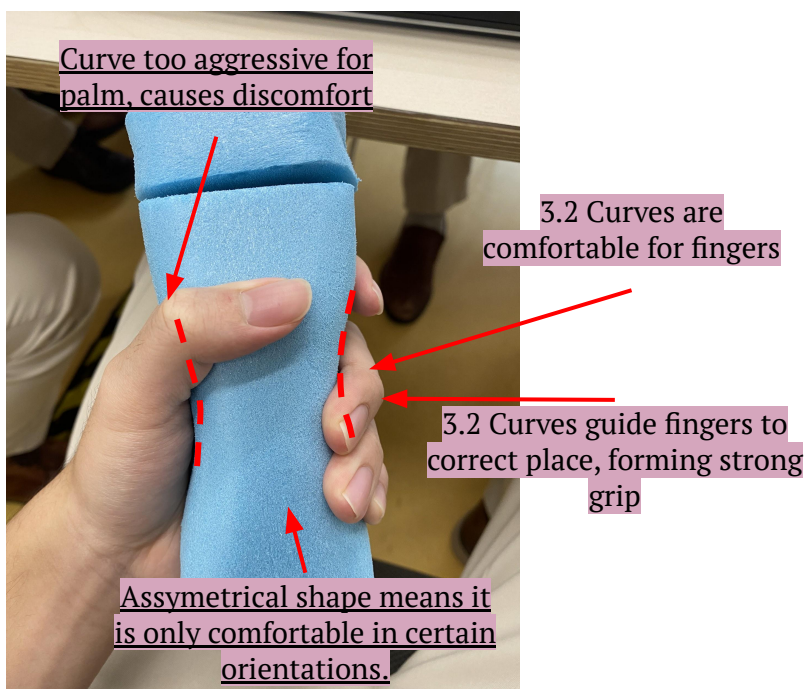


4.5 Design #3

Function: 90%
 Aesthetics: 90%
 Safety: 80%
 Ergonomics: 80%
Limitations



(foam model testing)



Judgement & Overall: 3/5

- Device is ergonomic and easy to handle, but as it is asymmetrical, it is uncomfortable in certain positions
- Grooves on side assist user in holding the device in a stronger position
- Lock mechanism is identifiable, but too similar aesthetically to cartridge
- Concavity on end of cartridge may make it more difficult to align to skin at a correct angle during use

Function: 90%
 Aesthetics: 90%
 Safety: 80%
 Ergonomics: 80%
 Limitations

4.6 Design #4

5.1 When orange stripes are aligned, the device is locked

1.1 Lock direction indicators

1.1 Top half is rotating lock

3.1 Uniform cylindrical body

Label

Transparent polycarbonate & glossy HIPs blend

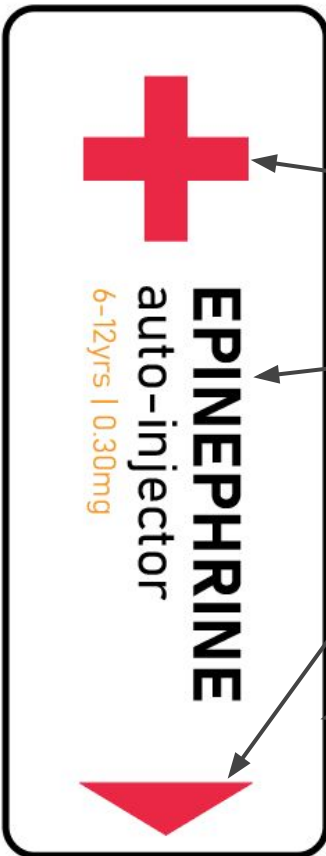
2.3 Replaceable cartridge containing syringe & solution

2.2 Mechanism activated by pushing orange end towards skin

Device may be difficult to hold when unlocked during use

7.1 Height: 150mm

7.1 Diameter: 35mm



5.1 Bold red colors for identifiability

5.1 font for identifiability

5.1 Red arrow indicating bottom-end of device

Insufficient space for instructions & use guidance

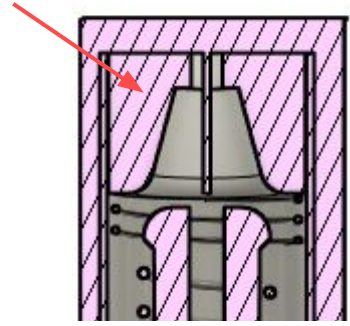
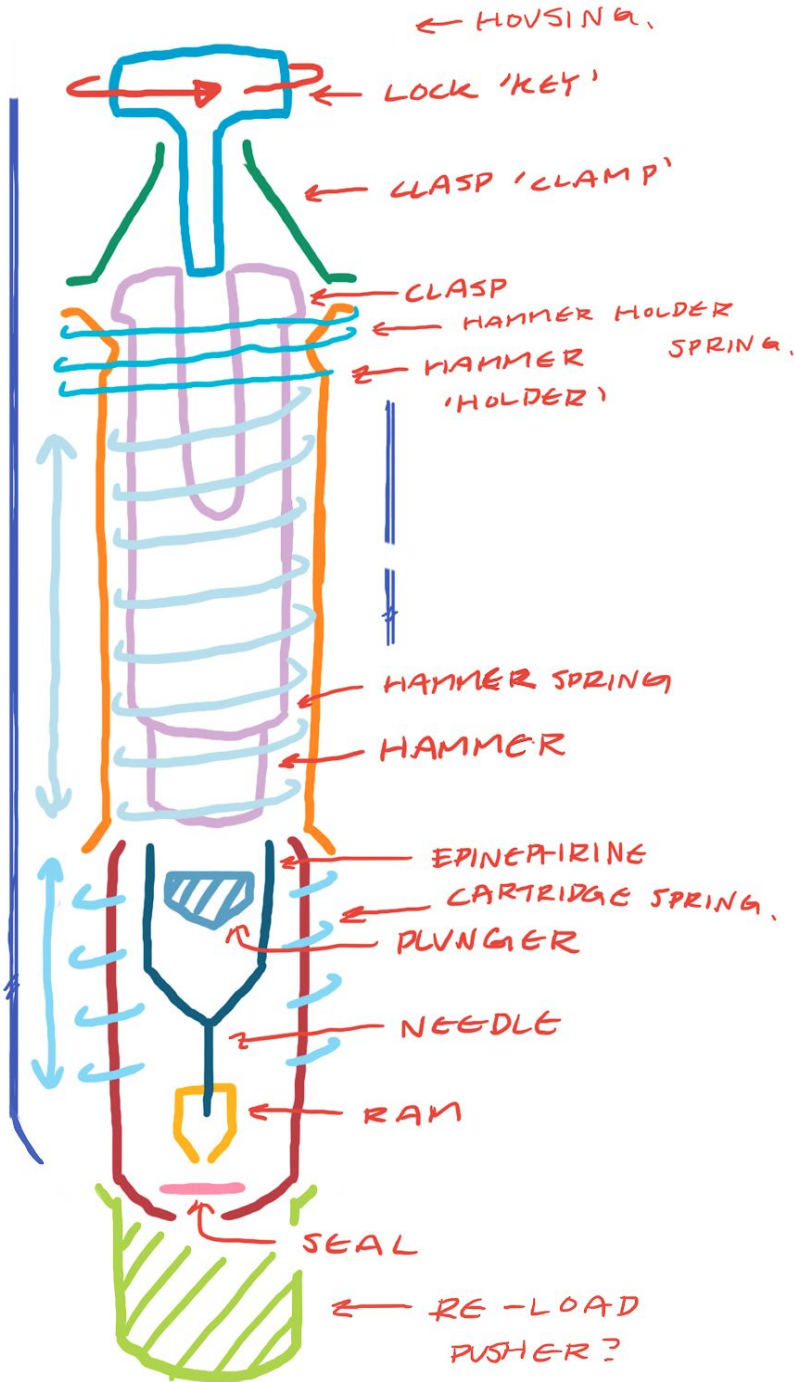
Judgement & Overall: 3/5

- As the lock rotates at the center of the device, when unlocked, it may be awkward to hold the device which prevents a strong fist-grip
- Insufficient space for instructions, reducing use guidance significantly
- Orange stripes help indicate when they device is locked or unlocked through visual and tactile feedback
- Transparent polycarbonate & glossy HIPs creates interesting and appealing aesthetic.

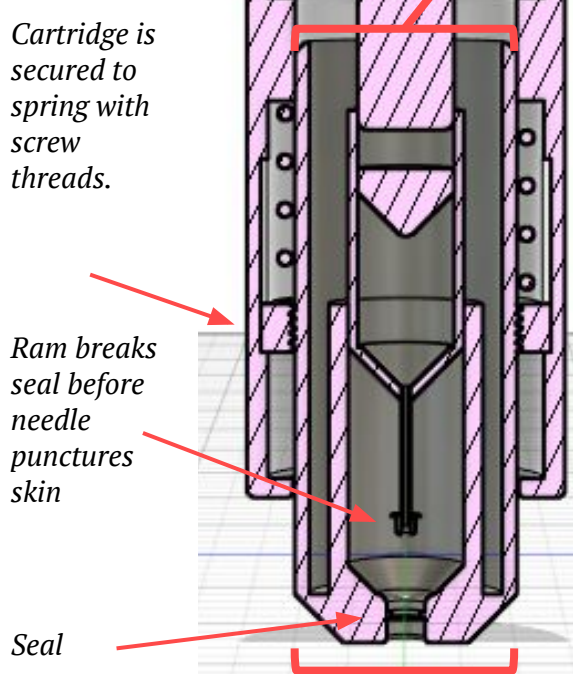
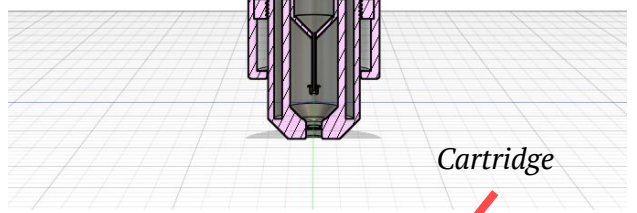
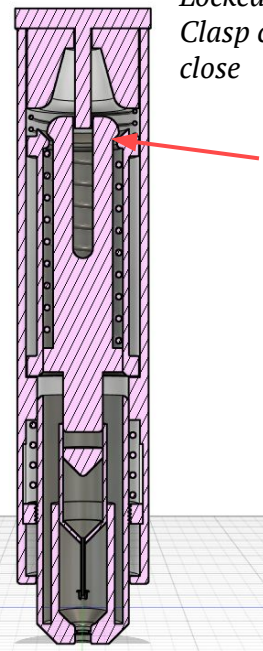
4.7 Conceptual Modelling

Mechanism Prototype

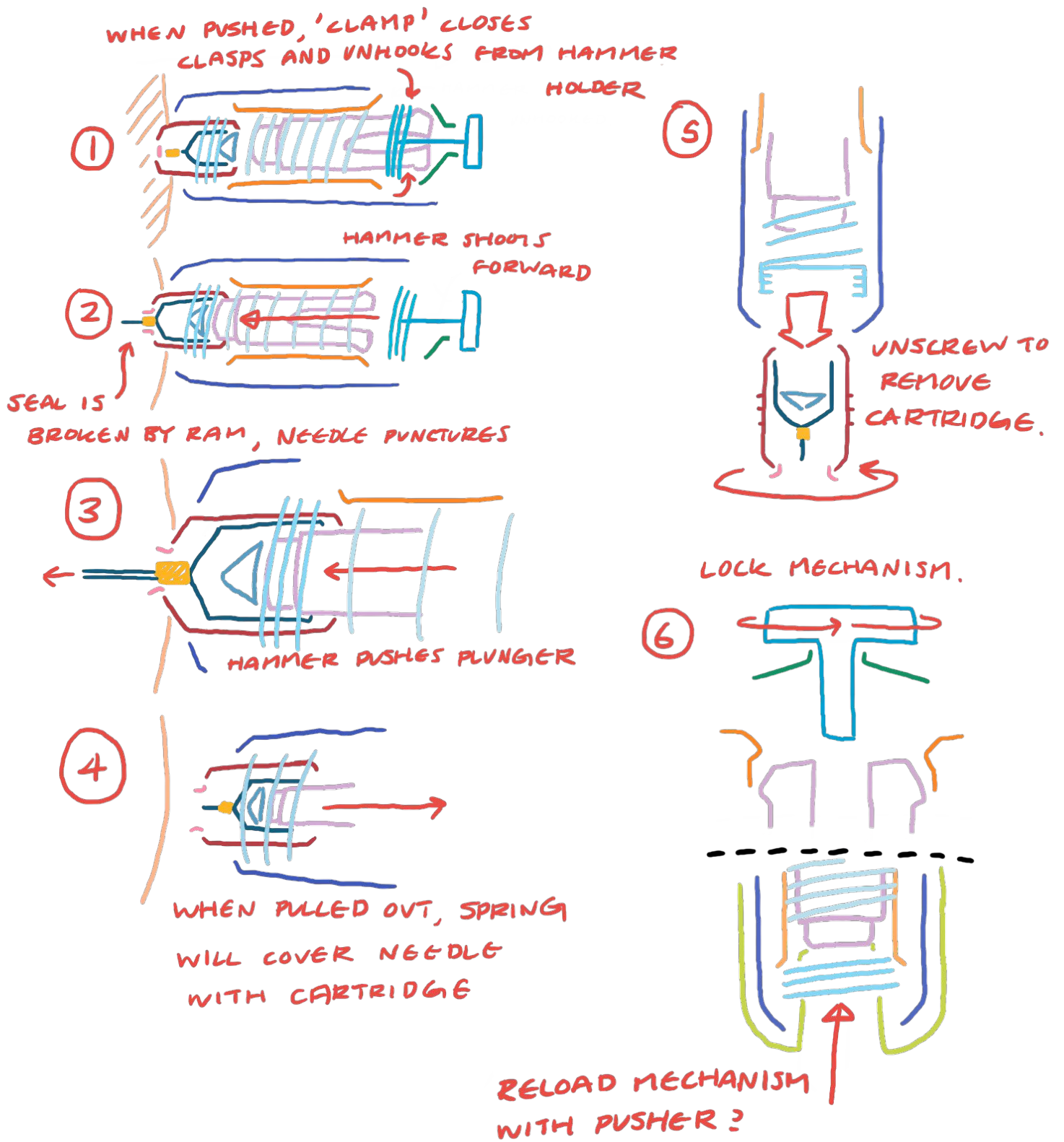
Unlocked state. Clasp is free to close.



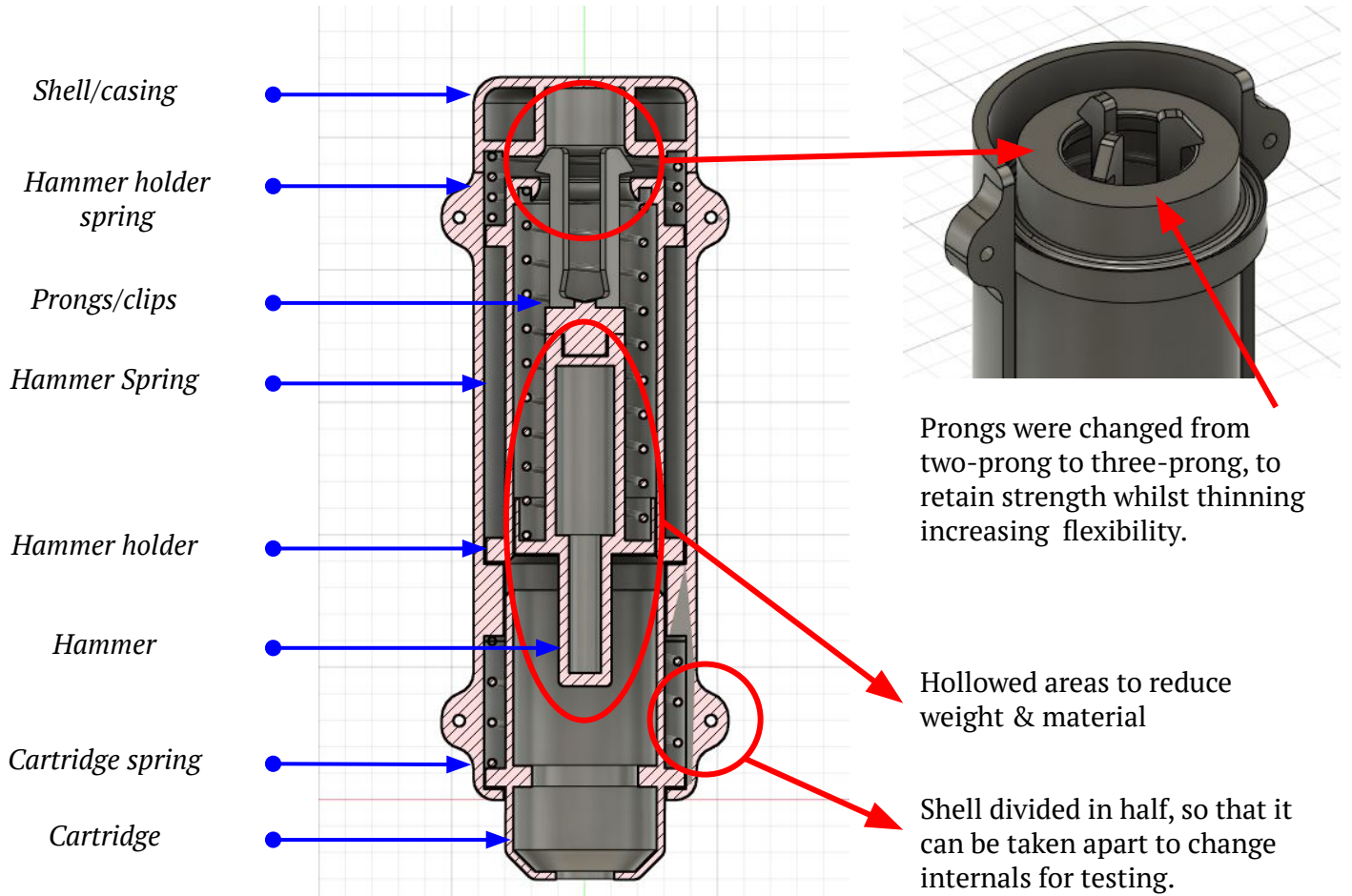
Locked state. Clasp cannot close



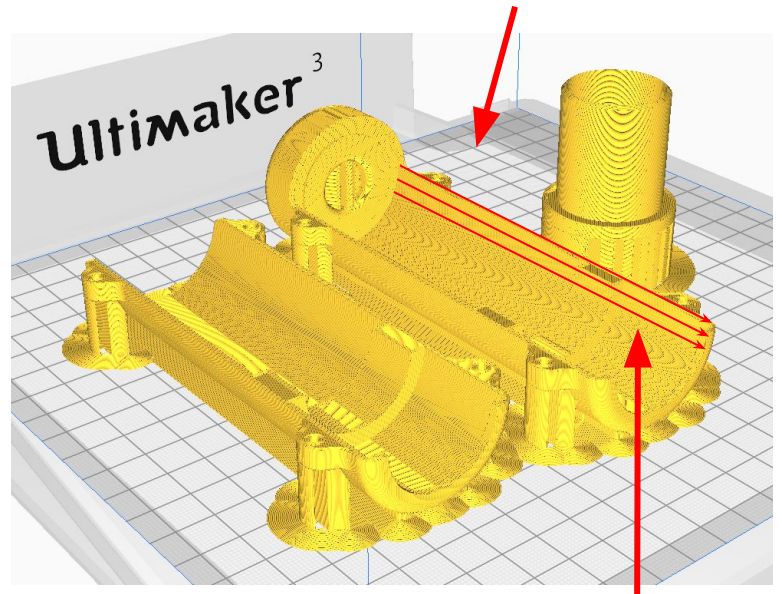
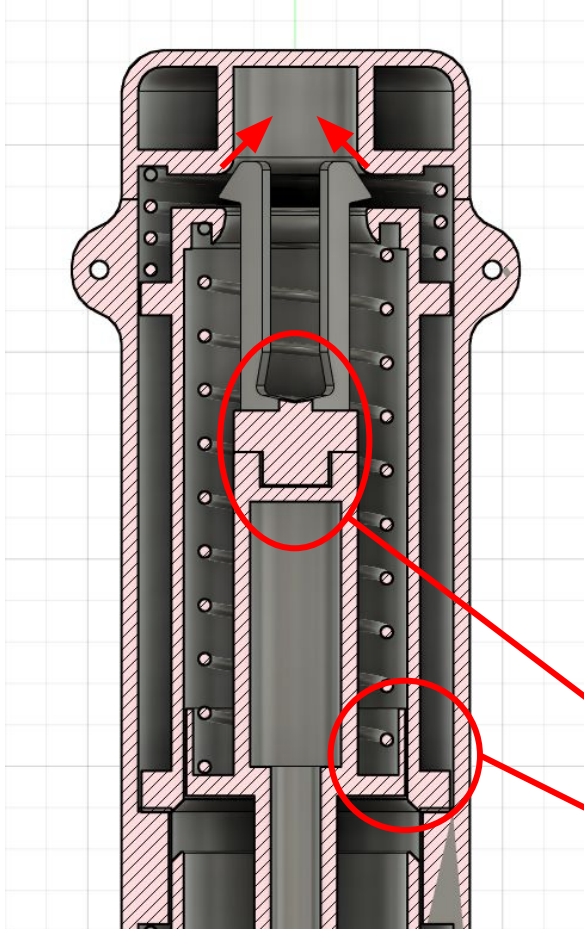
Mechanism Drawing



Iterated mechanism from initial design:



Prongs clasp inwards to unhook from hammer holder, firing the hammer.



Shell printed lengthwise to retain strength under tension from springs.

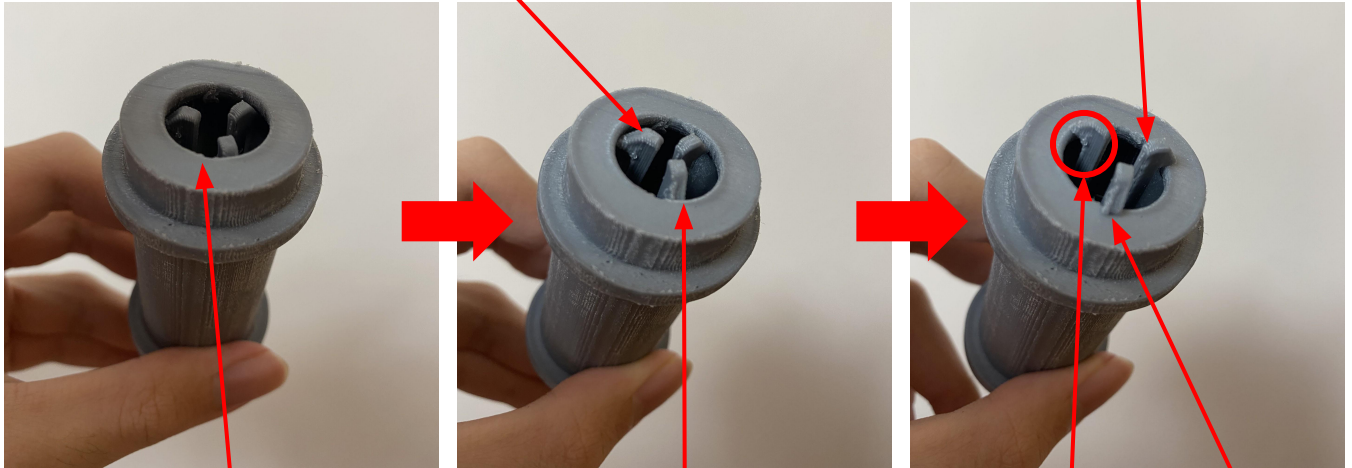
Prongs printed separately from hammer to replace if broken.

'Cone' added around hammer to hold spring.

Prongs/hammer mechanism

Prongs do not break after repeated deformation ✓

'Snaps' for audio & tactile feedback ✓



Prongs before being pushed through

Prongs are flexible enough to collapse and fit through hole ✓

Angled tops on prongs allow it to be pushed through ✓

Prongs prevent spring from ejecting ✓

Trigger / Activation Mechanism

Neutral position

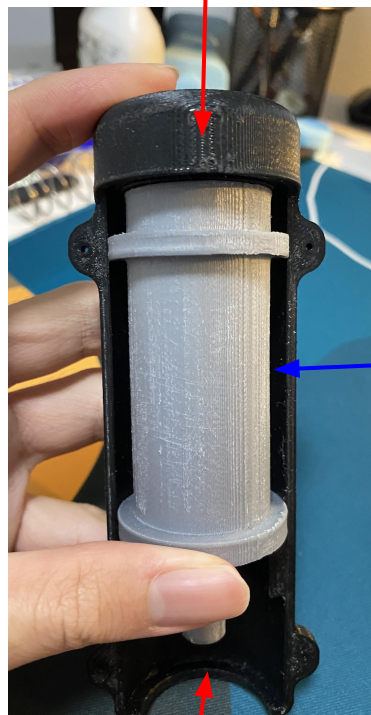
Prongs are successfully dislodged when hammer holder is pushed upwards ✓

Hammer & hammer holder

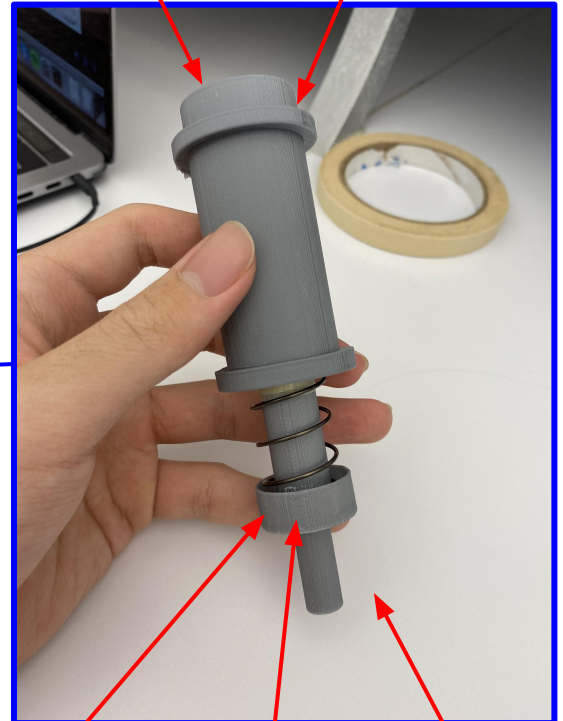
Spring inside hammer holder to eject hammer



Too much friction between holder & case ✗



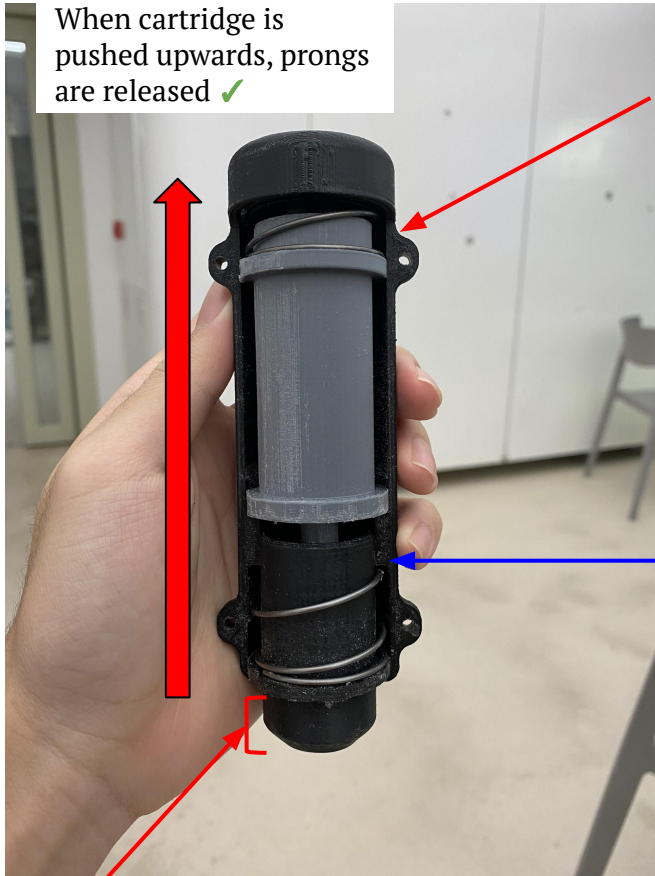
Hammer is released ✓



'Cone' to hold spring

Spring ejects hammer rapidly ✓

Parts printed lengthwise helps distribute force from spring ✓

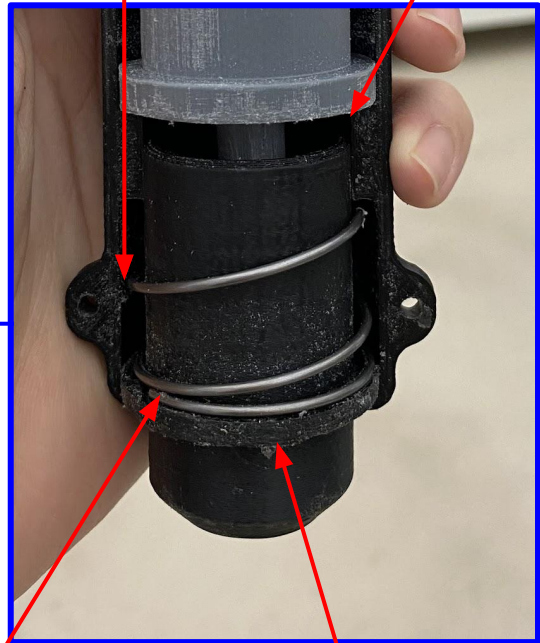


When cartridge is pushed upwards, prongs are released ✓

Spring to keep hammer holder in downwards position ✓

Chamfered ledge to push into hammer holder

Spring to keep capsule in downwards position ✓



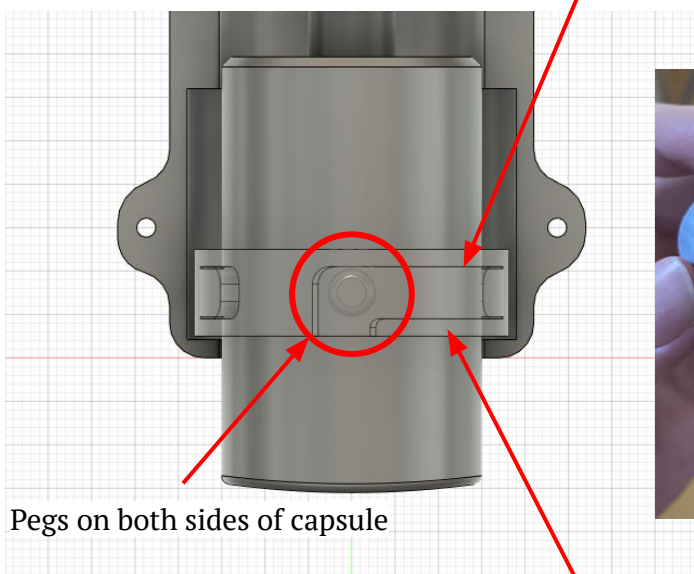
~14mm of travel distance, may be too short ✗

Spring does not provide enough force, too easy to push ✗

Too much friction between cartridge & case ✗

Reloadable cartridge (Iterated mechanism from initial design):

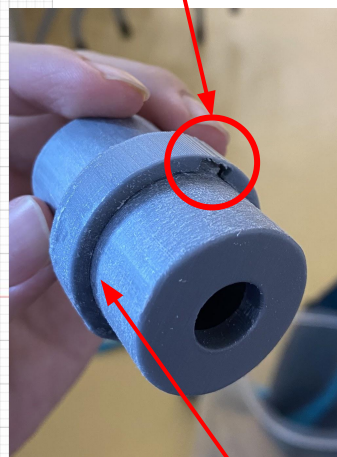
Twist-lock mechanism to secure capsule



Pegs on both sides of capsule

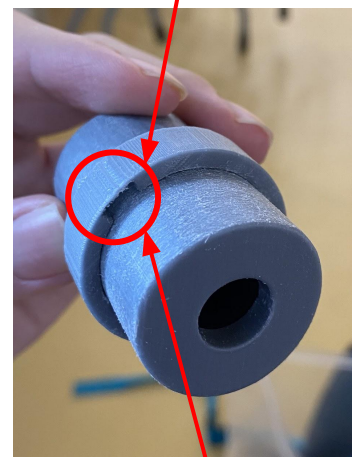
Cartridge holder is not secured to shell ✗

Unlocked position (pegs exposed)



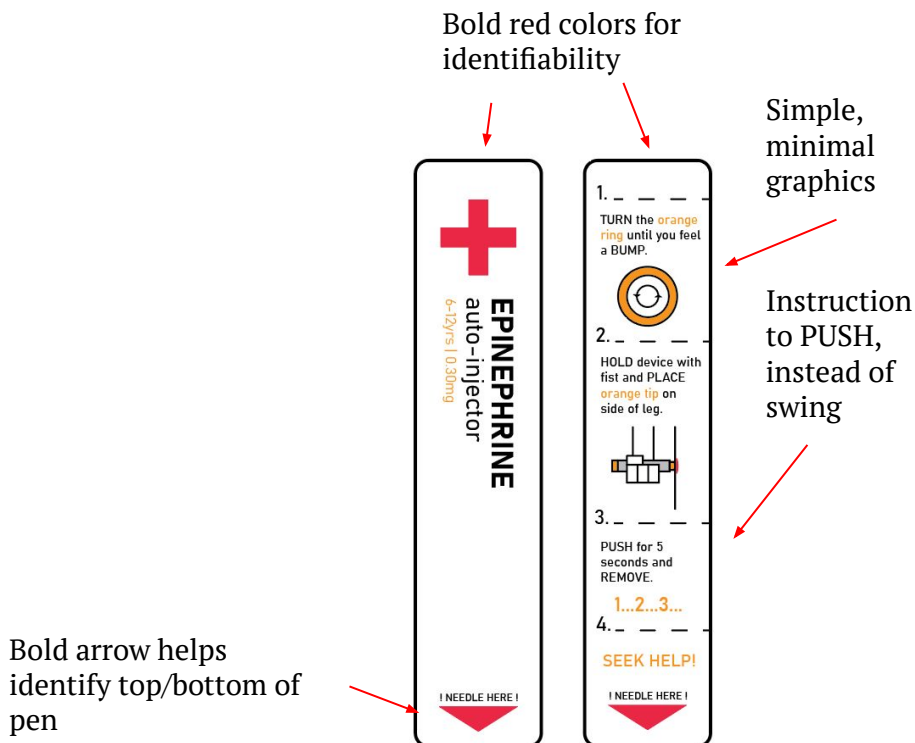
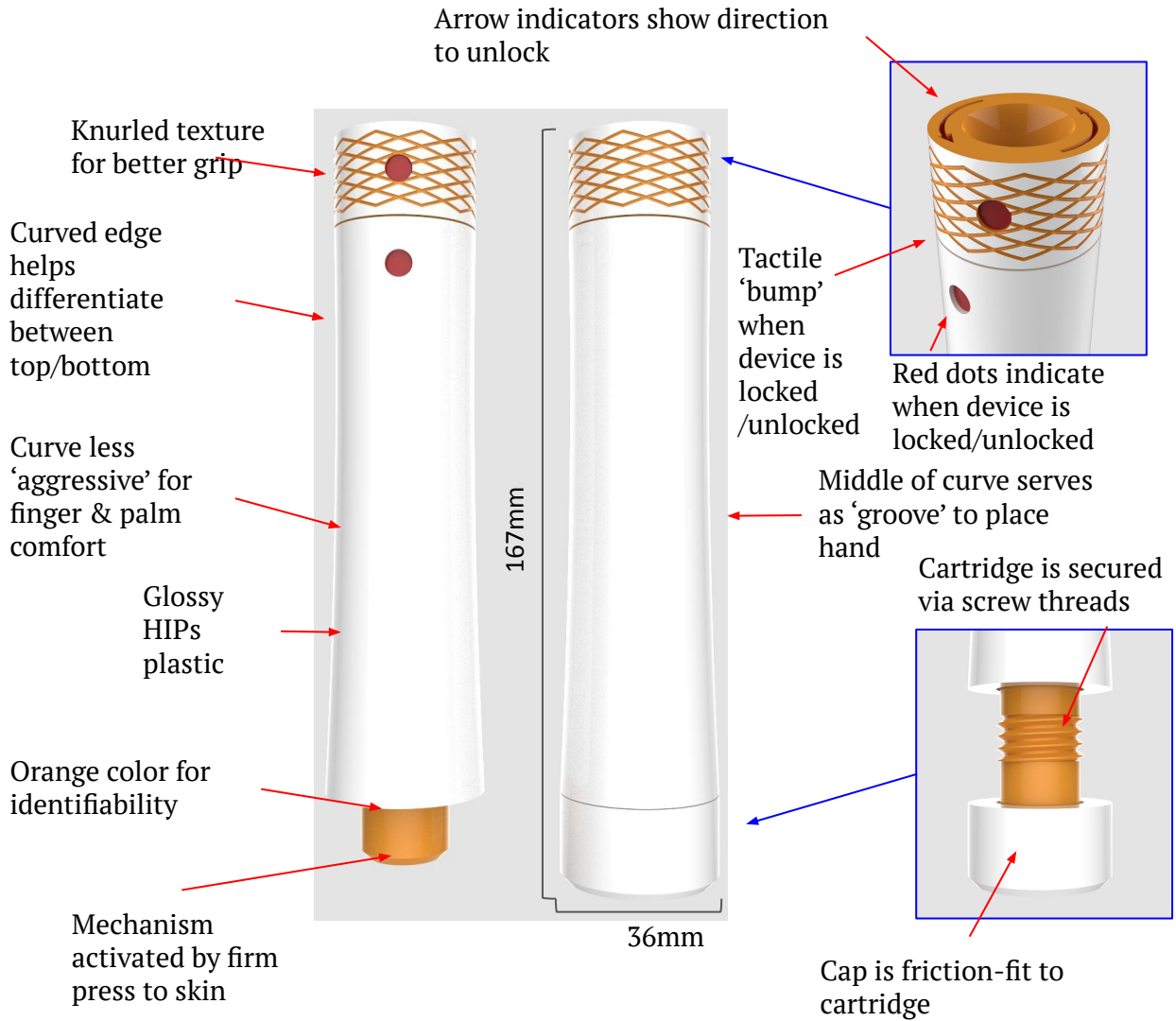
Too much friction between holder & case ✗

Locked position (pegs enclosed)



Poor indication of lock/unlocked state ✗

5 Final Design



Specification	✓/~/✗
1.1 Lock that prevents misfires	✓
1.3 Graphical use guidance & affordance	✓
2.2 Mechanism activated rapidly and intentionally	✓
2.3 Doses of epinephrine are replaceable	✓
3.1 Price range of device 200-250USD. Each cartridge priced 40-50 USD.	(testing)
4.2 Mostly cylindrical/symmetrical housing	✓
5.1 Height 150mm, diameter ~30-35mm	✓
6.1 Children aged 6-12. Poorer family backgrounds, suffer from anaphylaxis	(testing)
7.2 Graphics are simple and easy to understand.	✓

Justification

Design #1 has been developed for the final design. Modifications have been made - the body is now made of PLA, allowing it to be constructed via FDM to reduce costs. The body is more uniform with a shallow curve, achieving affordance whilst allowing it to be held in different orientations comfortably by the user. This also allows more space for graphics, to increase clarity of instructions and use guidance. The lock has a knurled texture to increase grip, and the red dots indicate when the device is locked. When twisted 180 degrees, a small tactile 'bump' can be

felt to indicate that the device has been locked. This design poses many advantages over conventional EAI's.

Mainly, the design fulfills the function of being reusable, increasing affordability and reducing costs to the consumer in the long-term. In addition, It better achieves use guidance through more accurate instructions, anthropometrics, and ergonomics. The device also has high compatibility to existing EAI's, as it activates through pressing the cartridge towards skin. This increases familiarity and reduces the learning curve.

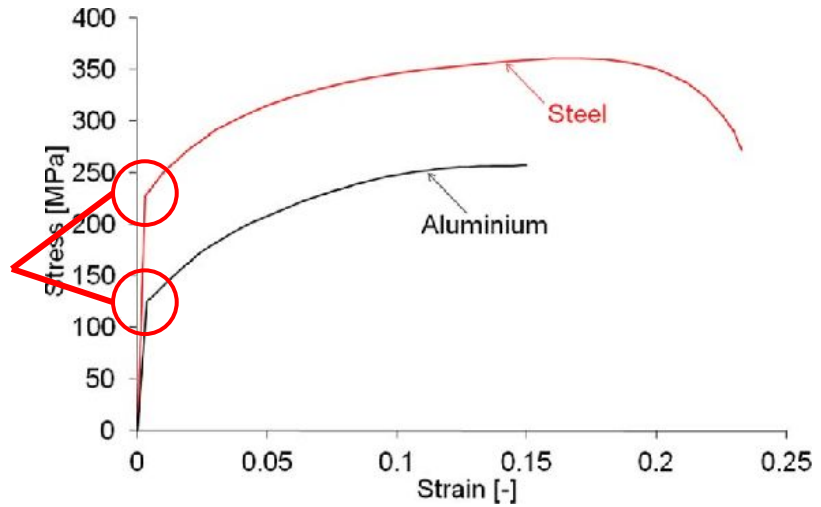
6 Materials and Processes

= select, = reject, xxx= important properties

Component	Material	Cost	Advantages	Disadvantages	
Spring	Carbon steel (alloy)	~0.70-1.80\$/kg	- very high yield strength/elasticity (Beck., 2020) - Very high hardness	- Lower ductility & malleability (difficult to manipulate into a coil) - lower corrosion resistance	
	Aluminium	~2.17/kg	- low density - very high ductility & malleability (AZoM, 2002)	- lower yield strength/elasticity - low hardness	

Point of stress at which material would be permanently deformed

Steel > aluminium
(Liu, et. al, 2013)



Hammer Prongs, shell, lock, cartridge, hammer, hammer holder, cartridge holder, cover/cap	PLA Plastic	0.05\$ per cm ³	<ul style="list-style-type: none"> - Cheap and easy to manipulate (via FDM) - Very low cost - biodegradable - High stiffness (Rogers, 2015) 	<ul style="list-style-type: none"> - Highly brittle - Very low thermal resistance 	
	ABS Plastic	0.04\$ per cm ³	<ul style="list-style-type: none"> - Very low cost - High toughness (COEX) - Low density - High thermal resistance 	<ul style="list-style-type: none"> - Difficult to recycle - More difficult to manipulate via FDM (due to high thermal resistance) 	
	Poly-carbonate Plastic	n/a	<ul style="list-style-type: none"> - Very high impact resistance (BPF) - Low density 	<ul style="list-style-type: none"> - Very low hardness - High thermal expansion 	

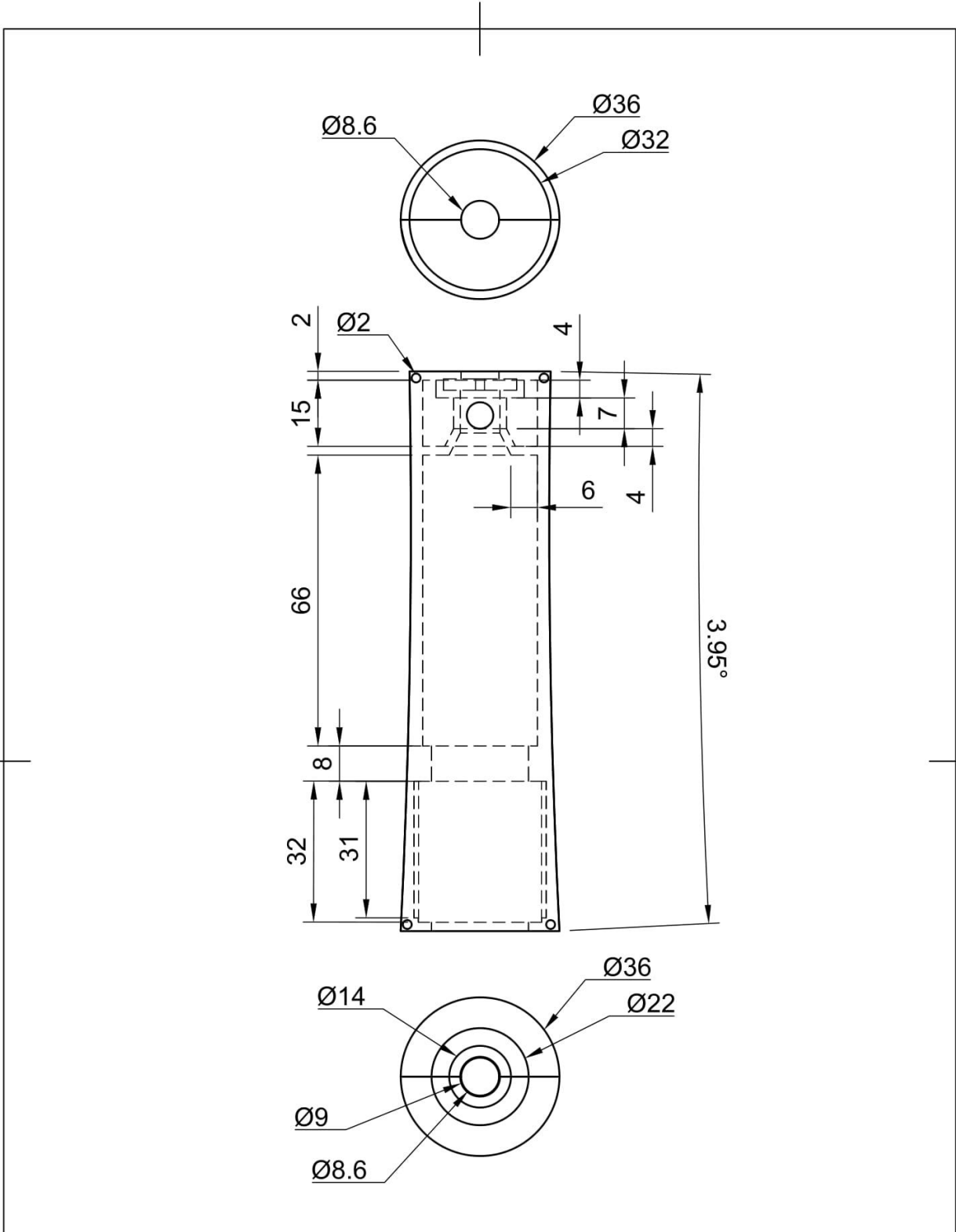
Label	Vinyl	0.99-1.75 \$/kg	<ul style="list-style-type: none"> - high hardness & corrosion resistance (AZoM, 2001) - Good printing ability 	<ul style="list-style-type: none"> - Only joining through adhesion - high cost 	
	Poly-propylene	n/a	<ul style="list-style-type: none"> - high hardness & corrosion resistance (BPF) - Allows for thermoforming 	<ul style="list-style-type: none"> - Lower printing ability/quality - susceptible to UV degradation overtime 	

■ = familiar, ■ = somewhat familiar, ■ = unfamiliar

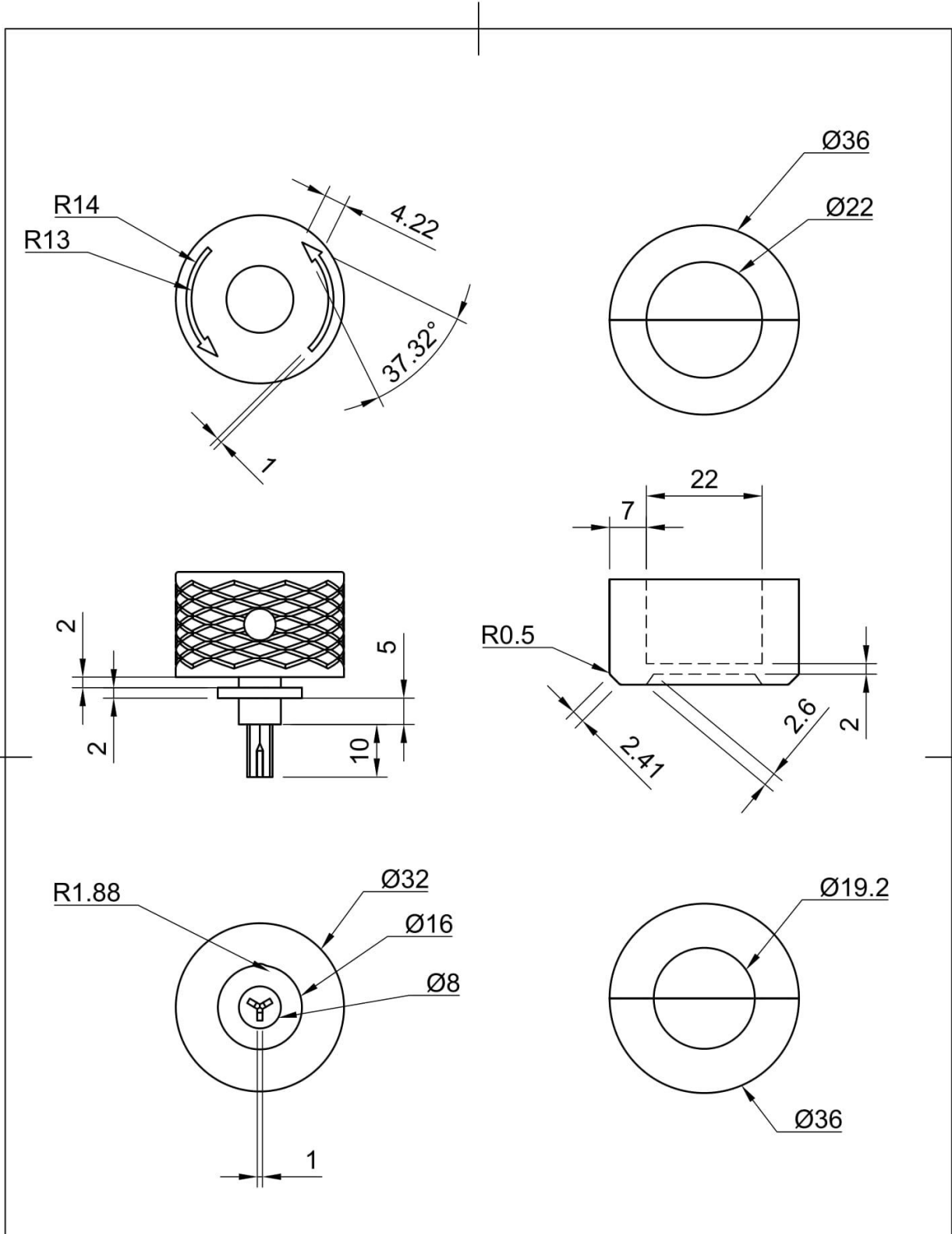
Method	Tools & Materials	Skill	Advantages	Disadvantages
Fused Deposition Modelling	3D-Printer Filament	■	- Allows for rapid prototyping with affordable materials - Enables the use of CAD	- Post-processing needed to address 'ripples' - 'Ripples' may cause lower durability
Stereo-Lithography	Resin printer, Filament	■	- Allows for rapid prototyping with high detail and quality - Enables the use of CAD	- High cost of filament (resin)
Spray Painting	Aerosol Paint	■	- Allows for rapid painting - High durability of paint applied - Application of different coatings or finishes	- Risk of uneven finish/error (lower margin for error)
Fasteners	Screws, Screwdriver	■	- Allows the prototype to be disassembled/reassembled when needed	- Temporary joining techniques may not provide sufficient durability overtime
Vinyl Stickers	Vinyl Sticker	■	- High durability and quality of print	- Adhesives may fail overtime
Heat-transfer	Heat-transfer Paper, Printer, Iron	■	- Adhesion for a longer period of time	- Lower quality and durability of print

Bill of materials					
Part	Volume (mm³)	Weight (g)	Quantity	Material	Cost (USD)
Shell	39820	42.2	1	PLA Plastic	\$0.88
Hammer (& prongs)	4813	5.10	1		\$0.20
Hammer holder	4640	4.92	1		\$0.10
Cartridge Holder	2504	2.66	1		\$0.10
Cartridge	5317	5.64	1		\$0.12
Cap	13390	14.2	1		\$0.30
Lock	9572	10.1	1		\$0.21
Spring (hammer holder)	234	1.84	1		Carbon Steel
Spring (hammer)	481	3.78	1	\$0.50	
Spring (cartridge)	464	3.644	1	\$0.50	
Screws	/	/	4	Stainless Steel	\$0.50

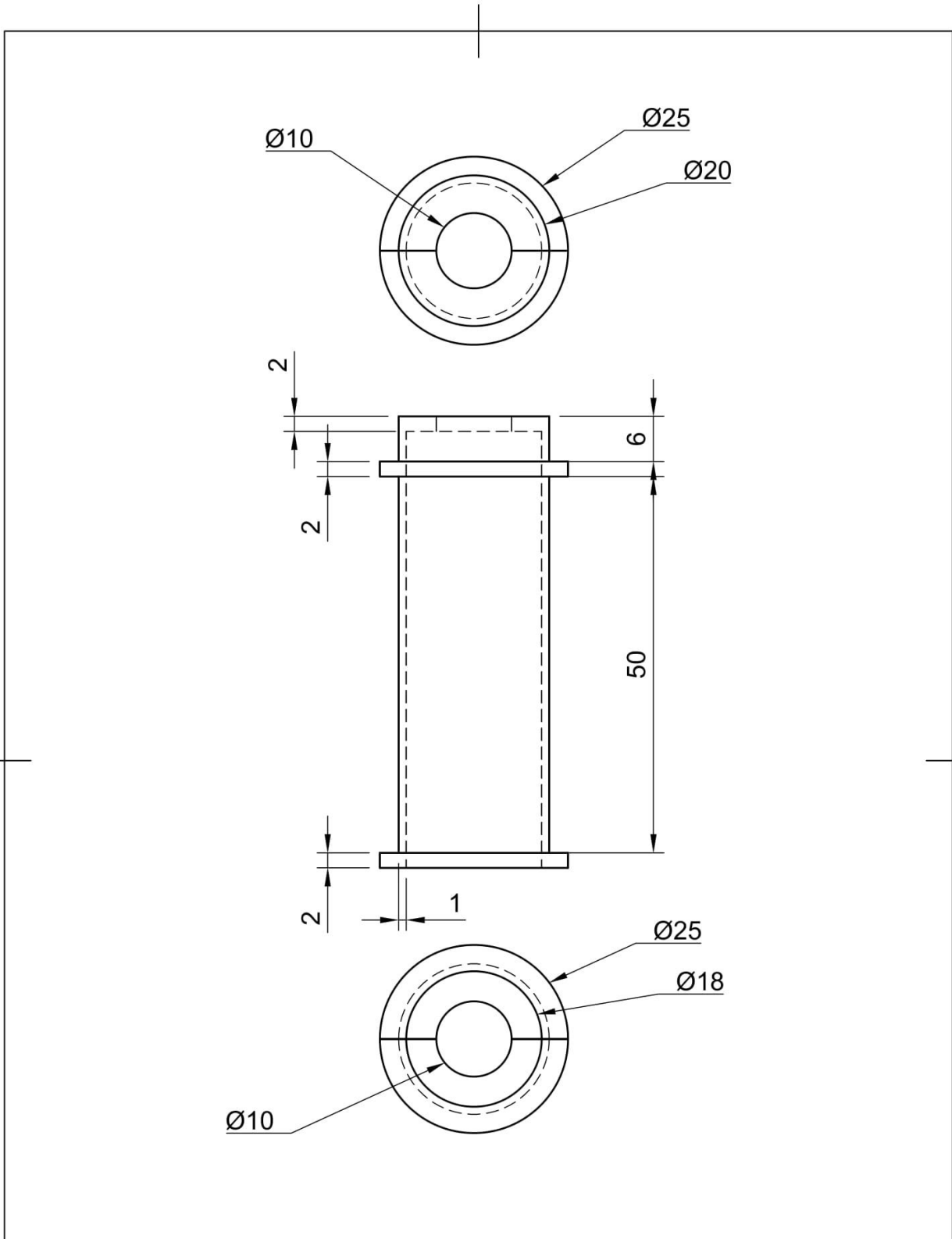
7 Orthographic Drawings



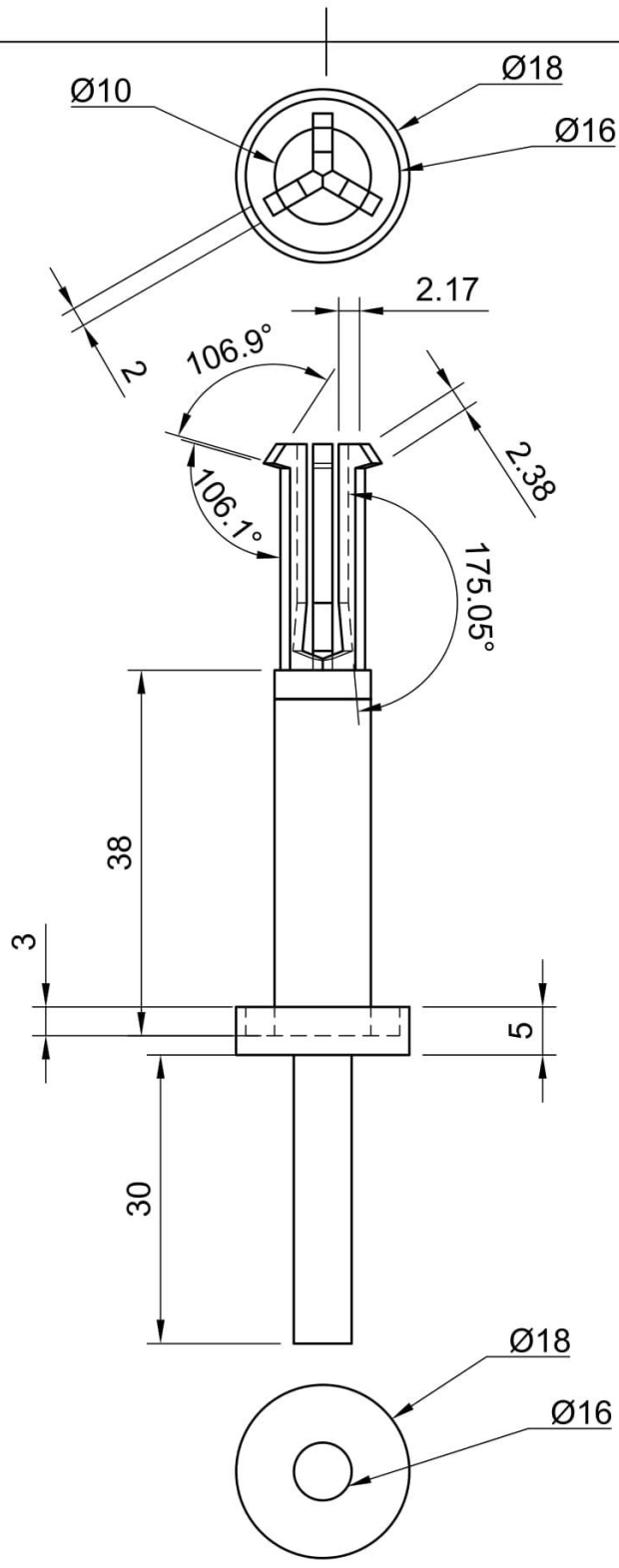
Dept.	Technical reference	Created by	Approved by	
		Document type	Document status	
		Title	DWG No.	
		EAI Shell	Rev.	Date of issue



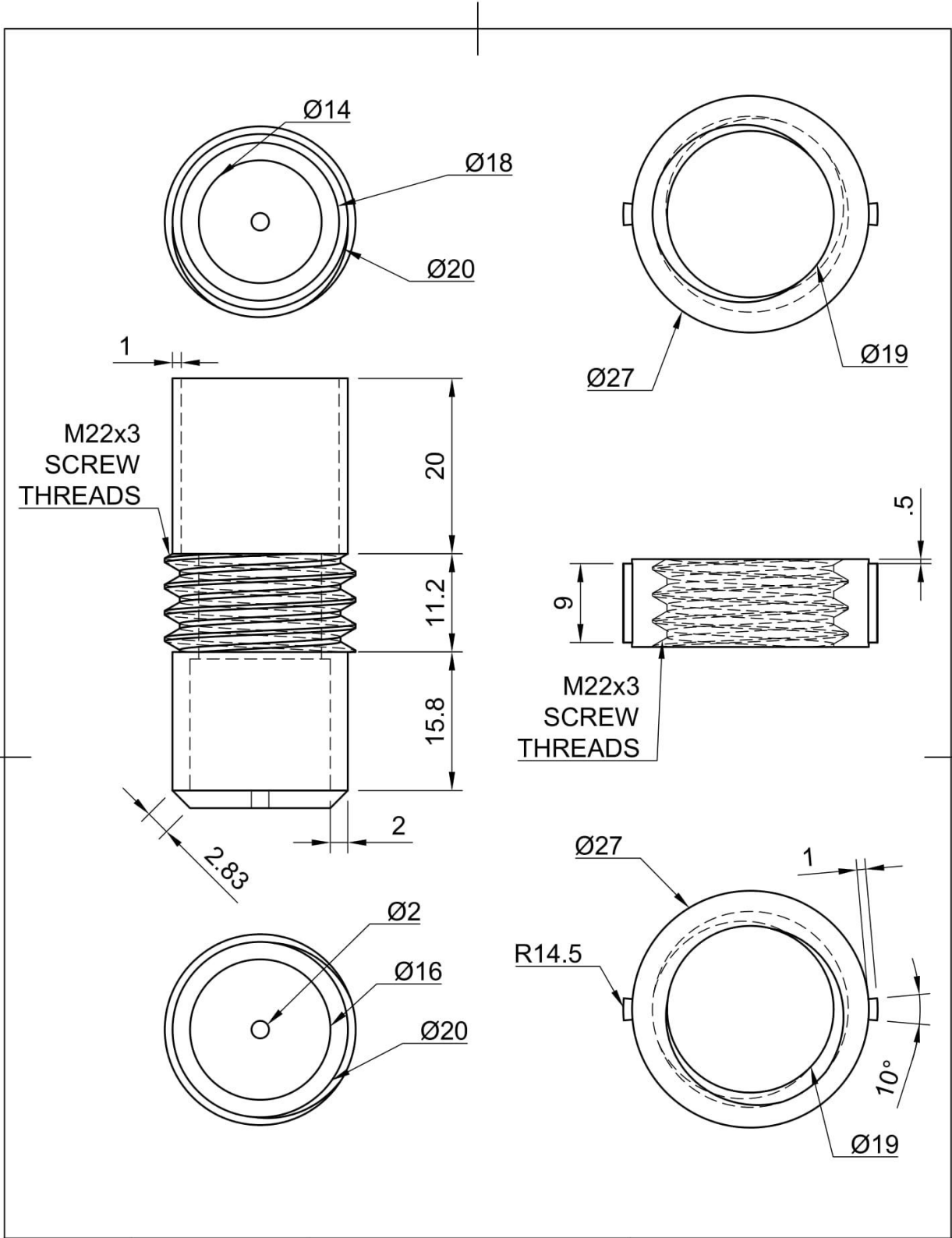
Dept.	Technical reference	Created by	Approved by	
		Document type	Document status	
		Title	DWG No.	
		EAI Lock (left) Cap (right)	Rev.	Date of issue
			Sheet 2/6	



Dept.	Technical reference	Created by	Approved by	
		Document type	Document status	
		Title	DWG No.	
		EAI Hammer Holder	Rev.	Date of issue

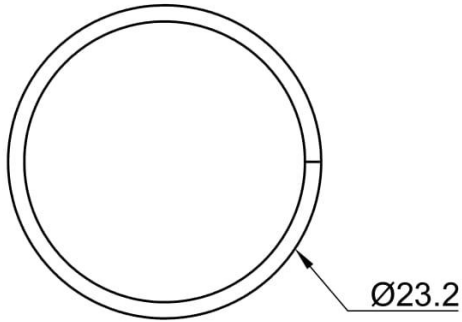


Dept.	Technical reference	Created by	Approved by		
		Document type	Document status		
		Title	DWG No.		
		EAI Hammer	Rev.	Date of issue	Sheet
				4/6	

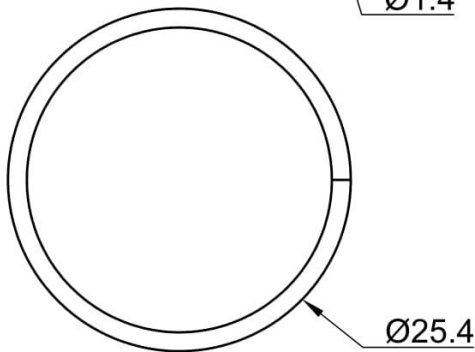
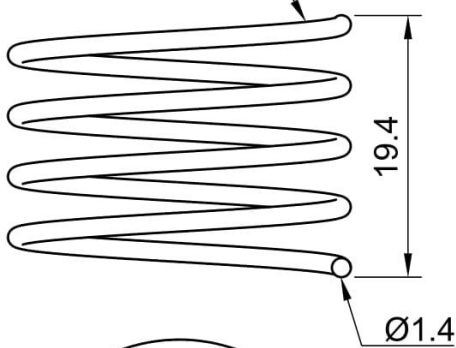


Dept.	Technical reference	Created by	Approved by		
		Document type	Document status		
		Title	DWG No.		
		EAI Cartridge (left) Cartridge Holder (right)	Rev.	Date of issue	Sheet 5/6

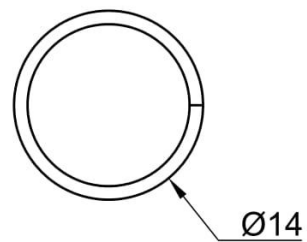
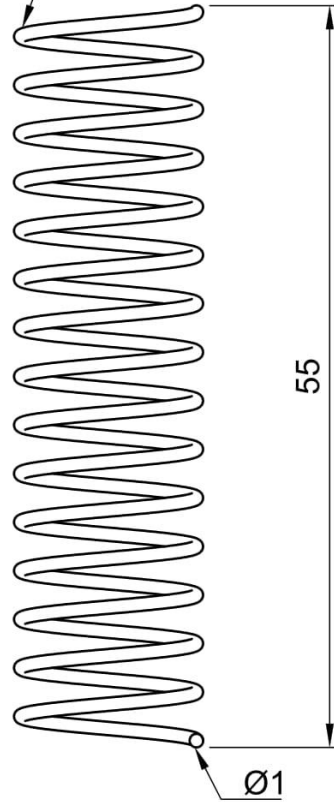
Hammer Holder Spring



Cartridge Spring



Hammer Spring

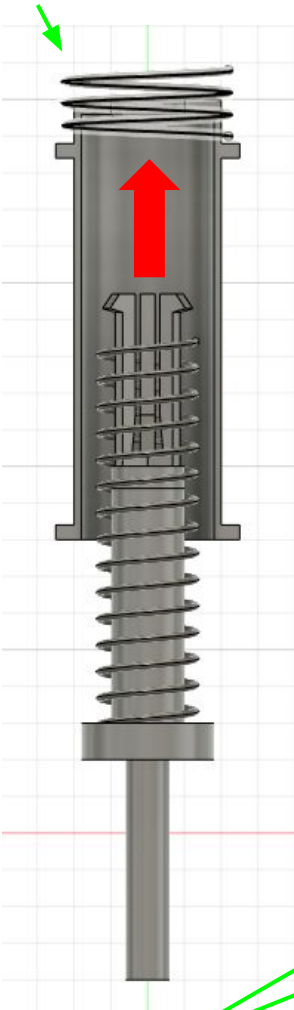


Dept.	Technical reference	Created by	Approved by		
		Document type	Document status		
		Title	DWG No.		
		EAI Springs		Rev.	Date of issue
				6/6	

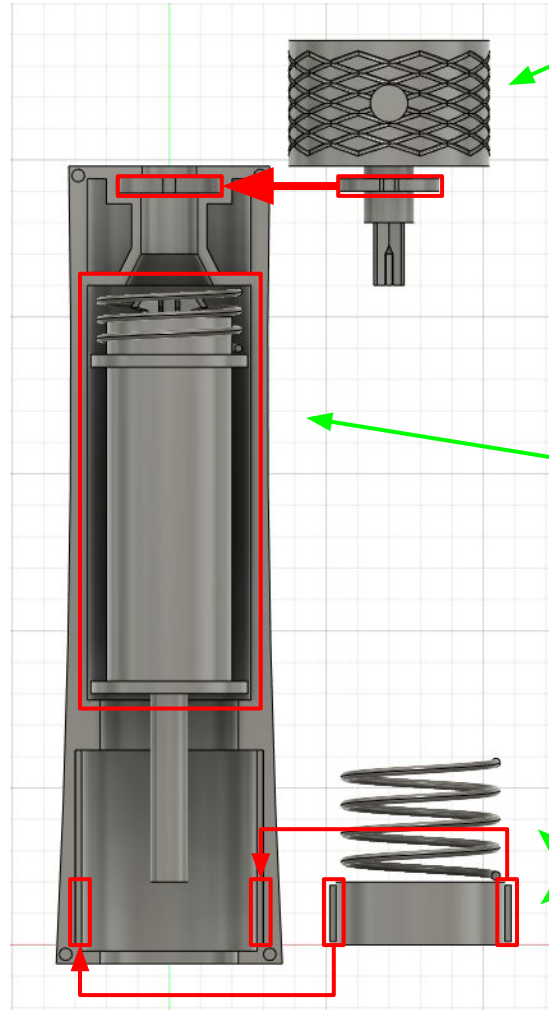
8 Plan of Production

Processes	Equipment	Time	Quality Assurance	Risk Assessment
Creating final prototype in CAD	Computer	8h	Purchase springs to ensure compatibility and correct dimensions	Sedentariness. Ensure sufficient rest in between work.
Export to print software & adjustments (print orientation, print setting, infil, quality)	Computer	15 mins	Ensure correct tolerance and adjust printer settings accordingly	n/a
Print via FDM (3D Printer)	Computer, 3D Printer, PLA, filament	24h	Ensure print plate is level and there is sufficient PLA filament.	Heat burns from hot plate/nozzle on 3D Print and warping. Allow print to fully cool before removing.
Post processing (support removal)	Pliers, File	30 mins	Do not rush process and cut support carefully.	Potential damage to print. Avoid by cutting support carefully and not forcefully.
Post processing (sanding) <i>*for components with tight/incorrect tolerances</i>	Sandpaper, file	45 mins	Do not rush process and sand surfaces evenly.	Potential damage to print. Avoid by sanding surfaces lightly & evenly.
Post processing (painting) <i>*for external components i.e. shell</i>	Spray paint	2h	Do not rush process (short bursts of paint, ensure enough time to dry).	Harmful fumes from aerosol paint. Work in a ventilated room/booth
Assembly of hammer (see diagram for assembly)	Springs	10 min	Ensure correct length of spring when cutting by referencing design proposal.	Deformation to spring when cutting. Avoid by using large pliers that can supply sufficient force.
Assembly of shell & other components (see diagram for assembly)	Screw-Driver, 3mm bolts, 33mm nuts	10 min	Ensure correct tolerances in screw holes. If not, return to sanding with a file or use a drill.	Cracking to shell due to tight fits. Avoid excessively screwing.

1. Insert the hammer spring and hammer into the hammer holder. Place the hammer holder spring on the top ledge of the hammer holder.

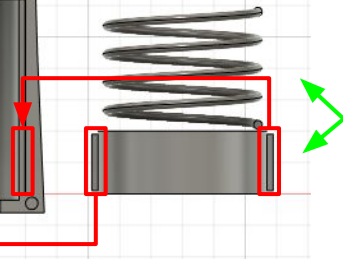


2. Align and place the lock within the top cavity of the shell as shown on the diagram.



3. Place the now assembled hammer holder and hammer within the middle cavity of the shell as shown.

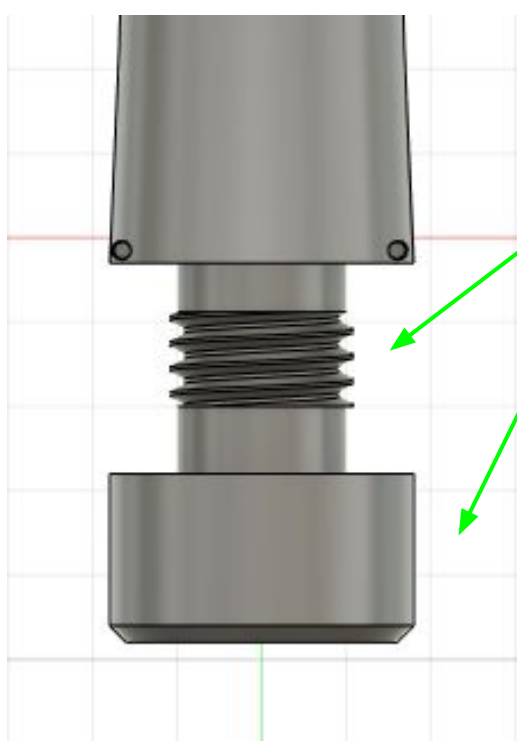
4. Place the remaining spring and cartridge holder within the bottom cavity of the shell. Align the 'wings' of the cartridge holder and place as shown on the diagram.



5. Insert screws to the areas as indicated and fasten.



6. Insert and fasten the cartridge and cap.



9 References

- 2002, Written by AZoMJun 4. "Aluminum - Advantages and Properties of Aluminum." AZoM.com, 13 Dec. 2019, www.azom.com/article.aspx?ArticleID=1446.
- "2019 Poverty Guidelines." ASPE, 22 May 2019, aspe.hhs.gov/2019-poverty-guidelines.
- "ABS Filament 1 3D Printer Filament 1 3D Materials 1 COEX LLC." COEX, LLC, 26 Oct. 2018, coexllc.com/about-abs-filament/.
- "Anaphylaxis." *Harvard Health*, Mar. 2019, www.health.harvard.edu/a_to_z/anaphylaxis-a-to-z#:~:text=Symptoms%20of%20anaphylaxis%20usually%20occur,is%20followed%20by%20vigorous%20exercise.
- Anshien, Marco, *et al.* "Unintentional Epinephrine Auto-Injector Injuries: A National Poison Center Observational Study." *American Journal of Therapeutics*, U.S. National Library of Medicine, 2019, pubmed.ncbi.nlm.nih.gov/27984266/.
- "Average Hand Size: Adults, Children, and Relation to Height." *Medical News Today*, MediLexicon International, 3 July 2020, www.medicalnewstoday.com/articles/average-hand-size#children.
- Beck, Kevin. "High Carbon Steel Properties & Uses." *Sciencing*, 16 Feb. 2020, sciencing.com/high-carbon-steel-properties-uses-7596348.html.
- Bpf. "British Plastics Federation." Polycarbonate (PC), www.bpf.co.uk/plastipedia/polymers/polycarbonate.aspx#:~:text=PROPERTIES,or%20special%20grades%20even%20lower.
- Brown, Allan. "Determining Your Power Grip Size." *MEMIC Safety Blog*, MEMIC, 30 Dec. 2014, memicsafety.typepad.com/memic_safety_blog/2014/12/determing-your-power-grip-size.html#:~:text=To%20measure%20your%20hand%20size,is%207.5%E2%80%9D%20multiply%20it%20by%20.
- "Developmental Milestones for Fine Motor Skills." Buffalo Hearing & Speech Center, askbhsc.org/developmental-milestones-for-fine-motor-skills.
- Kid Sense Web. "Fine Motor Development Checklist." *Kid Sense Child Development*, Kid Sense Child Development, 29 Nov. 2016, childdevelopment.com.au/resources/child-development-charts/fine-motor-developmental-checklist/.
- Anshien, Marco, *et al.* "Unintentional Epinephrine Auto-Injector Injuries: A National Poison Center Observational Study." *American Journal of Therapeutics*, U.S. National Library of Medicine, 2019, www.ncbi.nlm.nih.gov/pubmed/27984266.
- "Epinephrine - Drug Summary." *PDR*, Prescriber's Digital Reference, 2019, www.pdr.net/drug-summary/Adrenalin-epinephrine-3036.
- Ghoreshi, Kayvon. "Redesigning and Incorporating the EpiPen into a Smartphone Case." *UCONN Library*, University of Connecticut, 2013.
- Hazardous Waste Listings*. EPA, Sept. 2012, www.epa.gov/sites/production/files/2016-01/documents/hw_listref_sep2012.pdf.
- "High Impact Polystyrene (HIPS)." *Custom Plastic Injection Molding & Extrusion Professionals - Houston, Texas (713)643-6577*, Blackwell Plastics, www.blackwellplastics.com/HIPS.html.
- "Steel Material Properties." *Www.steelconstruction.info*, www.steelconstruction.info/Steel_material_properties.
- Herman, Bob. "EpiPen's Evolving Market SHare." *Axios*, Axios Media, 28 Mar. 2019, www.axios.com/epipen-mylan-market-share-generic-978d621c-4d0b-4f2a-bf64-d3ef25988039.html.
- Liu, Bin & Villavicencio, Richard & Guedes Soares, Carlos. (2013). Failure characteristics of strength-equivalent aluminium and steel plates in impact conditions. 10.1201/b15120-25. Manuyakorn, Wiparat, *et al.* Needle Length for Epinephrine Prefilled Syringes in Children and Adolescents: Is One Inch Needle Appropriate? *Asian Pacific Journal of Allergy and Immunology*, apjai-journal.org/wp-content/uploads/2018/06/8-AP-020317-0039.pdf.

- McCray, Chirlane. "EpiPens Save Lives. If They Cost Too Much, Kids Will Die." *The Washington Post*, WP Company, 1 Mar. 2019, www.washingtonpost.com/posteverything/wp/2016/08/25/epipens-save-lives-if-they-cost-too-much-kids-will-die/.
- Pancare, Rachel. "How Do Bright Colors Appeal to Kids?" *Sciencing*, 2 Mar. 2019, www.sciencing.com/do-bright-colors-appeal-kids-5476948.html.
- "Preventable Injuries From Life-Saving Epinephrine Auto-Injectors On The Rise." *ACAAI Public Website*, 3 Nov. 2015, www.acaai.org/news/preventable-injuries-life-saving-epinephrine-auto-injectors-rise.
- Rogers, Tony. "Everything You Need To Know About Polylactic Acid (PLA)," www.creativemechanisms.com/blog/learn-about-poly-lactic-acid-pla-prototypes.
- Rudden, Jennifer. "Average Rent Affordable for Low-Income U.S. Households 2019." *Statista*, 29 Oct. 2019, www.statista.com/statistics/1064468/average-rent-affordable-for-low-income-households-usa/.
- Tirrell, Meg, and Dan Mangan. "FDA Says There Have Been 'Hundreds of Complaints' about EpiPen Misfires, Some Followed by Patient Deaths." *CNBC*, CNBC, 7 Sept. 2017, www.cnn.com/2017/09/07/fda-says-there-have-been-hundreds-of-complaints-about-epipen-misfires.html.
- Turner, Paul J, *et al.* "Fatal Anaphylaxis: Mortality Rate and Risk Factors." *The Journal of Allergy and Clinical Immunology. In Practice*, Elsevier Inc, 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5589409/.

Hybrid Sentiment Analysis system to extract Language Bias in news media

Sally Sijie Song 宋思婕

Abstract

The modern news media is often biased. By pushing a certain perspective through its news narrative, biased news outlets are able to use their influence to manipulate public perception of information. Amongst other methods of swaying their audience, strategic word choices are frequently incorporated in news reports. This method is usually successful, the primary reason being the audience is unaware of how most news articles are biased. It is, therefore, the aim of this study to detect and quantify the affective states in the language used in modern news articles using a novel sentiment analysis model.

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool specially adapted to analyze short social media texts. In this paper, an error analysis of VADER is performed in a transductive transfer learning scenario; the pre-trained tool is then optimized using a Multinomial Naive Bayes classifier trained on movie reviews for detecting and quantifying sentiments in news articles. This is done manually after analyzing the outputs and combining the strengths of each model to create an improved hybrid system. This novel model that gives insight into the severity of bias in major global news outlets for later analysis, is presented and evaluated.

1 Introduction

The present-day news media exerts great influence over public opinion on the issues that are being reported. This power is often abused when news reports stray from objective facts and instead attempt to promote certain viewpoints over others.

Media bias can be accomplished through multiple means, such as the omission of information (intentionally omitting information that may discredit one's point) or the emphasis on certain views over others (e.g. including more liberal views than conservative views). It is difficult, however, to quantify information bias as there is no clear measure for the level of information transparency in journalism. Therefore tone bias, as it is relatively more detectable, usually functions as an indicator of bias for readers (Pannucci & Wilkins, 2011).

This paper defines biased language in news media as strategically selecting words that carry nonobjective attitudes with the intent to induce an affective response in the reader. Humans are often prone to persuasion by biased language, mainly

because news outlets tend to use only subtly biased language whose subjectivity is undetectable. Computer algorithms, on the other hand, use a fixed criterion which makes them capable of detecting and evaluating bias accurately and objectively.

Sentiment analysis (also referred to as opinion mining) is a natural language processing (NLP) task that uses data mining and text analysis to extract subjective attitudes from a body of text. A common use of sentiment analysis is extracting sentiments from social media content to recognize public opinion on certain topics or products.

2 Background

2.1 Historical Background

Sentiment analysis is one of the most rapidly developing fields of computer science. To understand the historical origin of sentiment analysis, Mäntylä. *et al.* (2018) used text mining and qualitative coding to analyze 6,996 papers from Scopus to generate a literature review on the evolu-

tion of sentiment analysis. They found that sentiment analysis was first utilized in text subjectivity analysis by the computational linguistics community in the 1990s, as well as for studying public opinion at the beginning of the 20th century. However, sentiment analysis research only gained notable popularity in 2004, as 99% of the papers from Scopus were from after 2004. Pang and Lee (2007) attributed the following factors as causes for this sudden increase in sentiment analysis research:

- the rise of machine learning methods in natural language processing and information retrieval;
- the increased supply of subjective texts on the World Wide Web, due to its popularity gain and, specifically, the exponential growth of online review sites;
- recognition of researchers for the wide range of applications and intellectual exploration that the area offers.

It was also found by Mäntylä. *et al.* (2018) that the usage of sentiment analysis has expanded significantly in recent years – the most studied sentiment analysis tasks have evolved from interpreting online product reviews for commercial purposes to analyzing texts on social platforms. The study of sentiment analysis has branched out further to connect with other areas of research, such as the stock market, elections, medicine, cyberbullying, software engineering, etc.

2.2 Sentiment Analysis

To define the characteristics of sentiment analysis, Pang and Lee (2007) identified its regression-like nature which differentiates it from other fact-based text-mining tasks such as part-of-speech tagging or named-entity recognition. They examined factors that make sentiment analysis a difficult task and concluded that the complex grammatical rules and devices of the English language means that the sentiments of a single word is fluid under different contexts, under situations where sentiment is expressed without the use of subjective words, the lexicon would be deemed ineffective. In conclusion, sentiment depends heavily on

the context. Thus it is difficult to quantify the sentiment expressed with a numerical unit of measure.

2.2.1 Sentiment Polarity

In polarity classification, where the task is to classify the text sentiment as either negative or positive, the line between subjective and objective information blurs. Pang and Lee (2007) specify examples where inherently objective information can convey subjective meaning based upon the context. For example, “battery life is 2 hours” is objective information but can be interpreted negatively when compared to the general standard for battery life. Some words, such as “democrats” can carry different sentiment polarity depending on the recipient of the information.

Pang and Lee (2007) also suggested other approaches that may be taken to create or assist a model for sentiment polarity classification. These approaches are summarized below:

- **Related categories:** analyzing text on other features related to polarity, such as categorizing reasons behind a certain sentiment (“I don’t like this computer because of its short battery life”) or ranking with comparative polarity (“I like this product less than the one last year”) or using outcome polarity such as recovery/death in medical journals to predict possible polarity.
- **Rating inference (ordinal regression):** analyzing data that provides the author’s self-rating on the sentiment expressed by their text (*i.e.* product reviews where users rate their satisfaction from one to five). This sort of analysis can arguably constitute a slightly different category – ordinal regression.
- **Agreement:** agreement detection (determining whether two pieces of text agree with each other)

2.2.2 Sentiment Intensity (Valence)

Yu and Hatzivassiloglou (2003) approached sentiment valence by classifying a text as either subjective, neutral, or objective.

Hatzivassiloglou and Wiebe (2000) established that adjective orientation has a prominent effect on the subjectivity of a sentence through studying the effects of dynamic adjectives, semantically oriented adjectives, and gradable adjectives on a subjectivity classifier. They created a novel method that statistically combines two indicators of gradability, showing that sets involving dynamic adjectives with positive or negative polarity or gradability are better predictors of subjective sentences than the class of adjectives as a whole.

2.2.3 Valence Aware Dictionary and sEntiment Reasoner (VADER)

Specially designed for the purpose of analyzing social media content, the sentiment analysis lexicon VADER is trained to evaluate short social media texts. This model, with its unique approach of employing pure manual human analysis in its lexicon generation, has the highest performance scores out of 11 state-of-the-practice sentiment analysis tools on four different domains and remains one of the most reliable sentiment analysis models. However, this model performed less successfully in other domains. Of the four domains tested, its performance score was the lowest on New York Times (NYT) Editorials, which also happens to be the domain with the longest text. Its overall precision, recall, and F_1 (the harmonic mean of Precision and Recall) scores on NYT Editorials is 0.69, 0.49, and 0.55 respectively, compared to its performance on a dataset of social media text (tweets), which yielded scores of 0.99, 0.94 and 0.94 respectively (Hutto & Gilbert, 2014).

Manually generated lexicons such as VADER are often rigid and carry drawbacks that, although insignificant, are difficult to resolve with manual approaches. In an attempt to resolve these errors and increase the accuracy of the VADER tool in the newswire domain, a probabilistic model – the Naive Bayes classifier – which generally displays more flexibility at domain adaptation than a rule-based model (Russell & Norvig, 2018), was combined with the VADER model.

Naive Bayes (NB) is a conditional probability classifier established on the basis of Bayes' theorem (its exponential form shown below):

$$P(y|X) = \frac{P(X|y) \cdot P(y)}{P(X)}$$

Where y and X are events, $P(y)$ is the probability of event y occurring, $P(X|y)$ is the probability of event X occurring given that event y has already occurred, and $P(y|X)$ is the probability of event y occurring given that event X has already occurred. The theorem calculates $P(y|X)$ under the naive assumption that all features in $P(y)$ and $P(X)$ are mutually independent. It then finds the maximum probability (y) through this equation where $P(X)$ is a constant:

$$y = \operatorname{argmax}_y [P(y) \cdot \prod_{i=1}^n P(x_i|y)]$$

VADER classifies input with regard to both polarity and valence. Created by Hutto and Gilbert (2014) in their research article “VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text,” it is rather unique in its approach: it is the only method in existing literature that completely strays from an ML approach by employing human labor to manually analyze text. It is also the only model reviewed here that detects both polarity and valence. Without a ML approach, this model is completely rule-based, combining grammatical rules with its human-generated lexicon. Below are some heuristic rules devised by Hutto and Gilbert (2014) to assess the polarity and valence of a sentence:

1. **Punctuation:** “I like this.” versus “I like this!”
2. **Capitalization:** “This place is amazing.” versus “THIS PLACE IS AMAZING.”
3. **Degree modifiers:** also called intensifiers, booster words, or degree adverbs. “Good” versus “Very good”
4. **“But”:** In the case of the contrastive conjunction “but”, the sentiment of the text following the conjunction is dominant. “I like the food here, but the service isn’t great.”
5. **Negation:** flips the polarity of the sentence 90% of the time, i.e. “The food here isn’t really all that great.”

The output of VADER is a number from -4 to 4 , with -4 representing an “extremely negative” sentiment and 4 representing an “extremely positive” sentiment. In this way, VADER measures the valence of sentiment on a regression scale.

Hutto and Gilbert (2014) then evaluated the VADER model against 11 typical benchmarks which include lexicons such as LIWC (Linguistic Inquiry and Word Count), ANEW (Affective Norms for English Words), the General Inquirer, SentiWordNet, and ML models relying on Naive Bayes, Maximum Entropy, and Support Vector Machine (SVM) algorithms. The results show VADER to have the highest classification precision, recall, and F1 scores compared to the other 11 benchmarks.

3 Method

In this paper, the VADER tool that is especially attuned to sentiments in social media was combined with a Multinomial Naive Bayes algorithm trained on tweets to create a novel hybrid system specialized in analyzing news articles for language bias. The model development was based on an analysis of the model in a transductive transfer learning scenario, where the performance of the VADER tool was first analyzed on a familiar domain, a dataset of NYT Editorials for error analysis, and was then tested on an unfamiliar domain, a dataset of recent online news articles. VADER’s performance was compared with the Naive Bayes model’s performance on the same dataset before the most successful characteristics of each individual model were combined and integrated.

The main challenge in finding training data for models in this paper is the lack of availability of sentiment-labelled newswire-based corpus (e.g. NYT Editorials) on the Internet. Thus, the ML model used in this paper is trained on domains of movie reviews or tweets.

The VADER tool was obtained from GitHub, titled under vaderSentiment by cjhutto¹. The lexicon was developed on a dataset of 4,000 tweets pulled from Twitter’s public timeline and 200 tweets

¹Link: <https://github.com/cjhutto/vaderSentiment>

specifically contrived to test syntactic and grammatical conventions.

The Naive Bayes model used the built-in classifier of Textblob, which was trained on the built-in movie review dataset in NLTK (the Natural Language Toolkit).

VADER’s ground truth (i.e. the “correct answer” the model uses as a target) on NYT Editorials as well as its performance were examined for error analysis (see Section 3.1), since opinion articles are heavily composed of rhetoric, difficult to analyze and therefore would induce an extensive amount of errors for analysis. This data was obtained from GitHub under the same directory as that of the VADER tool.

VADER was then tested on an unfamiliar domain, a manually collected small dataset of news articles on recent significant political events that underwent extensive news coverage. Only political events were chosen because other genres of factual news inherently contain polarity (i.e. climate change is bad, stock market surge is good, etc.); political news of topics such as terrorism that inherently contain strong polarity were also not chosen. The dataset consists of five snippets that were taken from each of the 30 political news articles, each article from one news outlet reporting on one selected political event.

Name of news outlet	Country of news outlet
BBC	United Kingdom
Reuters	
The Guardian	
CNN	The United States of America
NYT	
Fox News	

Table 1. Six influential English-based news outlets on which the model’s performance will be evaluated on.

3.1 Error Analysis on VADER

VADER’s performance on NYT editorials was significantly lower than its performance on social media text. To discover the cause for this, an er-

Event	Time of the event	Parties involved
Black Lives Matter protest	May 2020 – present	U.S., New Zealand, France, the U.K., Ireland, Brazil, Germany, Italy, Poland, Denmark, The Netherlands, Israel, Sudan
Donald Trump’s Impeachment	December 2019 – February 2020	U.S.
2019–20 Hong Kong protest	June 2019 – present	Hong Kong, Mainland China, U.S. (alleged), Taiwan
US–China Trade War	January 2018 – present	U.S., China
Brexit	June 2016 – January 2020	The U.K., the E.U.

Table 2. Five recent significant political events in the dataset.

ror analysis was performed on the ground truth of the model. VADER’s ground truth was the average human rating of 20 human raters collected in an anonymous survey. The ground truth of the NYT editorials was provided by Hutto and Gilbert (2014). Below are the characteristics observed from these human ratings of texts:

1. Some neutral, factual sentences were rated as having a sentiment intensity. For example, “On Dec. 19, around noon, New York City’s Union Square Business Improvement District plans the First Santa Claus Ulympics (that’s yule-lympics), in which about 50 contestants, in red suits and beards, will compete.” was rated as having a 0.4 positive sentiment. Some questions, holding no sentiment in itself, such as “Have countries threatened with retaliatory tariffs under the new trade law entered into negotiations with the United States?” were rated as moderately negative (−0.55), although the question is neutral, slightly negative at most.
2. On the other hand, some sentences that carry sentiment were rated as neutral. The sentence “It’s an economic euphemism with precedent,” carries a negative sentiment but was rated as neutral. This may be because the word “euphemism” is a noun that is generally neutral and its sentiment carried in this context of criticism or mockery wasn’t detected.

3. Sarcasm was not detected. For example, when rating the sentences “Some budget terms are technical necessities, like the distinction between actual outlays and budget authority. Others are gems of political euphemism,” both sentences were rated as having positive sentiments (0.05 and 0.95, respectively) even though the second sentence actually expresses criticism. This might be because “gems” and “euphemism” are relatively positive words. The sentence “No one has devised a euphemism of equal elegance for the other side of the ledger – until now,” was also rated as 1, extremely positive, even though it was a sarcastic sentence that expresses a negative connotation. There were many cases of subtle sarcasm in the NYT editorials that were not detected by the algorithm.

4. Consecutive sentences were not analyzed in accordance with each other, although it is often necessary to do so in order to analyze sentiments. For example, the sentences “The new Reagan budget is skimpy on detail and limp on initiatives, but page 2–13 offers a new pearl of circumlocution. The Administration, it says, calls for no cuts in benefit levels for several programs,” were given separate sentiment scores, −0.75 and 0.35 respectively. This analysis is incorrect because these two sentences support each other, thus they convey the same sentiment, which is a negative sentiment of relatively high intensity. A correct analysis of one of these sentences cannot be made without the context of the other sentence.

In NYT editorials containing 500 snippets of news articles, approximately five sentences out of the 10 sentences in the second snippet of the NYT Editorial had their sentiment polarity inappropriately labeled, as shown in table 3.

Since the ground truth for NYT editorials contains a noteworthy amount of inaccuracies, the accuracy report of VADER by Hutto and Gilbert (2014) in their research article may be incorrect. An error analysis on VADER’s performance on NYT editorials is therefore performed again on the same 10 sentences, as shown in table 4.

Sequence	Sentence	Polarity Score
1	Government budget time means budget language time.	-0.09
2	Some budget terms are technical necessities, like the distinction between actual outlays and budget authority.	0.01
3	Others are gems of political euphemism.	0.23
4	The recent prize-winner, which may be the work of an unknown New York state budget writer a few years ago, is revenue enhancement, meaning tax increase.	-0.05
5	No one has devised a euphemism of equal elegance for the other side of the ledger -- until now.	0.25
6	The new Reagan budget is skimpy on detail and limp on initiatives, but page 2-13 offers a new pearl of circumlocution.	-0.19
7	The Administration, it says, calls for no cuts in benefit levels for several programs.	0.08
8	For some others, however, the Administration proposes carefully targeted reforms.	0.03
9	Perfect: a description for spending cuts, even harsh ones, that sounds innocent, progressive, humane.	0.30
10	Is there, somewhere, a government budget in which carefully targeted reforms involve spending more money instead of less?	-0.11

Table 3. A 10-sentenced example of error spotting in sentences from the ground truth on NYT editorials with sentiment scores rated by 20 human raters. Ratings have been processed to range from -1 (extremely negative) to 1 (extremely positive). Correct ratings are highlighted in green, incorrect ones in red.

Sequence	Sentence	Polarity Score
1	Government budget time means budget language time.	0.00
2	Some budget terms are technical necessities, like the distinction between actual outlays and budget authority.	0.10
3	Others are gems of political euphemism.	0.00
4	The recent prize-winner, which may be the work of an unknown New York state budget writer a few years ago, is revenue enhancement, meaning tax increase.	0.08
5	No one has devised a euphemism of equal elegance for the other side of the ledger -- until now.	0.06
6	The new Reagan budget is skimpy on detail and limp on initiatives, but page 2-13 offers a new pearl of circumlocution.	0.00
7	The Administration, it says, calls for no cuts in benefit levels for several programs.	0.15
8	For some others, however, the Administration proposes carefully targeted reforms.	0.03
9	Perfect: a description for spending cuts, even harsh ones, that sounds innocent, progressive, humane.	0.06
10	Is there, somewhere, a government budget in which carefully targeted reforms involve spending more money instead of less?	0.03

Table 4. A 10-sentenced example of error spotting in sentences from a dataset of NYT editorials with sentiment (polarity and subjectivity) scores rated by VADER. Ratings have been processed to range from -1 (extremely negative) to 1 (extremely positive). Correct ratings are highlighted in green, incorrect ones in red

Table 4 shows that VADER labeled the polarity of two out of the ten example sentences correctly,

demonstrating a much lower accuracy than that of the ground truth. This remains true throughout all

of the snippets in the NYT Editorials dataset.

Through the error analysis performed on a dataset of snippets in NYT Editorials, it can be observed that the VADER lexicon itself has four main shortcomings:

1. **Lack of vocabulary:** The VADER lexicon has a relatively extensive vocabulary, yet there are words that are not labelled with sentiment in the lexicon—usually ones that are uncommonly used or formal, since the lexicon is designed to analyze short, casual social media texts containing mostly vernacular language. Therefore, quite frequently, VADER fails to label the sentiment of a certain word, “abomination,” for example, that doesn’t exist in the lexicon.
2. **General lack of context:** Phrases that are related to historical or cultural contexts, such as “The Statue of Liberty,” “concentration camp,” or “Garden of Eden” aren’t detected, as well as idioms such as “big headed,” even though some of these phrases are often used in an attempt to express a sentiment.
3. **Lack of inter-textual context:** VADER considers each sentence individually without the context of other sentences. Therefore, sentences that refer to each other, such as “The Democrats are optimistic. Other parties, however, feel the opposite way,” are always given an incorrect score.
4. **Lack of understanding of rhetorical devices:** VADER’s lack of inter-textual context leads to its inability to detect rhetorical devices. The detection of rhetorical devices is heavily dependent on the style of a text (e.g. humorous irony is expected to be seen in a text with a consistently humorous tone), which cannot be extracted without a comprehensive analysis of the whole text. Without inter-textual context, one cannot tell whether a sentence like, “Well that’s just great” is sarcasm -- it refers back to the last sentence.

The drawbacks of the VADER system listed above were unavoidable in a strictly rule-based system as it is incredibly difficult to manually design heuristics that encompass all of the subtleties involved

in human languages. It is therefore proposed to implement a probabilistic ML model to assist the tool, as its flexibility could potentially amend for flaws in the VADER system.

4 Results

4.1 Hybrid Model

Along with VADER, a multinomial Naive Bayes Classifier (see Section 2.2.3) was used in this paper to analyze five snippets that were taken from each of the 45 political news articles. The comparison of their outputs is shown in Table 5.

The polarity output from the Naive Bayes (NB) model was separated into two scores for positive and negative sentiments. These two scores were combined, allowing the positive and negative scores to negate each other (e.g. if positive score is 0.6 and negative score is 0.4, the sentence would be positive with an intensity of 0.2), for a compound score of the overall sentiment polarity in the sentence.

The outputs of the news dataset showed that the VADER model classifies polarity at a relatively high accuracy rate – much higher than the Naive Bayes classifier. The accuracy improved significantly compared to its analysis on NYT Editorials since news articles traditionally do not use as many rhetoric devices that may be difficult for VADER to detect.

A notable result is that in almost every article snippet, the Naive Bayes classifier is generally more proficient at scoring the subjectivity intensity of the text in the cases where sentiment intensity isn’t neutral. An example of this can be seen in the snippet analyzed in Table 5.

Therefore, to optimize the performance of both models, the polarity classification function of the VADER lexicon – the superior function – is combined with the subjectivity intensity detection of the Naive Bayes classifier.

5 Discussions

Through manual analysis of the novel model’s performance on the news articles dataset (see Section

Sentence	Rating		Ground Truth
	NB	VADER	
One year ago on Tuesday, hundreds of thousands of demonstrators in Hong Kong gathered for a march that became the start of the semi autonomous Chinese city’s biggest political crisis and the broadest expression of public anger with Beijing in decades.	0.99	-0.83	Medium–highly intense negative
In the months that followed, protesters filled the city’s streets, broke into the local legislature and vandalized it, staged sit-ins at the airport, and turned a university campus into a fiery battleground.	0.65	-0.83	Medium–highly intense negative
Earlier this year, the demonstrations quieted amid the coronavirus pandemic.	-0.33	0.00	Neutral
But Beijing’s push to impose national security laws over the territory has prompted some protesters to return to the streets.	0.63	-0.26	Slightly negative
Their presence is a reminder that many thorny issues – including the demonstrators’ demands for greater official accountability – remain unresolved.	0.99	0.10	Slightly positive

Table 5. An example of an article snippet. Five sentences chosen from the article being analyzed by the model. “Hong Kong Protests, One Year Later” from NYT (Ramzy et al., 2020). Ratings range from -1 (extremely negative) to 1 (extremely positive). Ground Truth is manually generated by myself.

3), it was concluded that the novel hybrid model is relatively more successful in quantifying bias and sentiment in news articles than each of the individual models tested.

VADER’s performance was accurate in its polarity mostly because it is a lexicon with non-objective words labelled with their specific polarity. On the other hand, the potential reason for the success of the Naive Bayes model in predicting non-neutral sentiment intensity is that a pattern usually exists in the sentiment intensity of a sentence: while sentiment polarity can easily be negated or flipped, the intensity for the sentiment often happens to be consistent throughout the text.

The Naive Bayes model’s performance could have been improved significantly if the training data belonged to the same domain as its testing data; having been trained on a dataset of movie reviews, due to the lack of availability of labeled news article datasets, the model’s function would have potentially been compromised. The ability to con-

sider intertextual context could have been a significant limitation, since movie reviews are at most a few sentences long and the correlation between each sentence did not hold as prominent an impact in terms of understanding the sentiment of the sentence.

Further work on sentiment analysis would include parsing techniques to improve inter-textual context. Parses would connect the subject of the sentences, therefore connecting sentences that refer to the same subject using determiners such as “this”. Furthermore, as different arguments and perspectives can be expressed in the same sentence, argument recognition can be used to detect different arguments and achieve a more nuanced analysis.

References

- Hatzivassiloglou, V., & Wiebe, J. M. (2000). Effects of adjective orientation and gradability on sentence subjectivity. *Proceedings of the 18th conference on Computational linguistics -*, 1, 299-305. doi: 10.3115/990820.990864
- Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8.
- Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27, 16-32. Retrieved 2021-04-13, from <https://arxiv.org/pdf/1612.01556> doi: 10.1016/j.cosrev.2017.10.002
- Pang, B., & Lee, L. (2007). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2, 1-135. <http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf> doi: 10.1561/1500000011
- Pannucci, C. J., & Wilkins, E. G. (2011). Identifying and avoiding bias in research. *Plastic and Reconstructive Surgery*, 126, 619-625. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2917255/> doi: 10.1097/prs.0b013e3181de24bc
- Ramzy, A., Ives, M., & Fei, L. Y. (2020). Hong Kong Protests, One Year Later. *The New York Times*. Retrieved 2021-04-13, from <https://www.nytimes.com/2020/06/09/world/asia/hong-kong-protests-one-year-later.html>
- Russell, S., & Norvig, P. (2018). *Artificial intelligence : A modern approach*. (4th ed.). Pearson Education.
- Yu, H., & Hatzivassiloglou, V. (2003). Towards answering opinion questions. *Proceedings of the 2003 conference on Empirical methods in natural language processing*, 129–136. Retrieved 2021-04-13, from <https://dl.acm.org/doi/10.3115/1119355.1119372> doi: 10.3115/1119355.1119372

Within areas of knowledge, how can we differentiate between change and progress?

Candace Yan Yue Chung 鍾欣瑜

The different areas of knowledge are never stagnant; there is always new knowledge to be discovered in every waking moment. However, whether the new knowledge gained is “change” or “progress” has always been a topic of dispute. This essay’s title suggests that in different areas of knowledge, change and progress are distinguishable. At first glance, the difference between “change” and “progress” seems obvious: change is a difference that can be realized immediately, while progress is a positive difference that, in retrospect, pushes you closer to the goal of the area of knowledge. But, is this always the case? Firstly, I will discuss how progress and change in both Natural Science and the Arts align with the established definition above. Then, I will argue that, ultimately, all change is a type of progress in the Natural Sciences, while all progress is a type of change in the Arts.

In the Natural Sciences, progress is a positive difference that, when viewed in retrospect, refine our understanding of the natural world, whereas change is just a difference that can be identified immediately. Initially, Fermi’s revolutionary paper on weak interactions - which explained beta radiation, a phenomenon that has remained elusive to scientists for years - was rejected by *Nature* in 1933 for being “too removed from reality”. The significance of Fermi’s paper was only recognized by the scientific community in 1938, earning Fermi the Nobel Prize (Macdonald). Since Fermi’s paper not only refine our understanding of beta radiation but also was only recognized in retrospect, his paper satisfies this essay’s definition of scientific progress.

Notably, the rejection of incorrect theories in the Natural Sciences also counts as progress as they

refine our understanding of the natural world. Scientific knowledge is acquired through inductive reasoning (making generalizations based on observational evidence), implying the validity of scientific theories is never certain, since there is always the possibility of contradictory new evidence appearing (Bradford). As such, scientific theories are more likely to be proven false than true. In 1811, Meckel postulated that human fetuses have “gill slits” because humans develop in stages corresponding to other “less perfect” species such as fish and reptiles (Vickers). However, Meckel’s theory was rejected after it was discovered that human fetuses have slits on their necks because of the genetic similarity between fish and humans instead (Vickers). Since this rejection reduces the number of viable explanations for why fetuses have slits on their necks, it indirectly refine our understanding of fetal development, thereby constituting progress.

In addition, progress spurred by science aims to produce something of utility. In 1939, Von Halban discovered that Uranium emitted approximately 3.5 (later corrected to 2.6) neutrons per emission, confirming the possibility of creating nuclear chain reactions (Von Halban *et al.*). The results of Von Halban’s experiments sparked the creation of technologies such as atomic bombs (weapons of mass destruction) and nuclear reactors (efficient clean energy generators), both of which are considered “useful” to modern society. Thus, as it sparks the production of useful innovations, Von Halban’s paper constitutes progress spurred by science.

Although examples of progress are easily identifiable in Natural Science, it is challenging to conceive what may constitute change since any

change that doesn't refine our understanding of the natural world or produce anything of utility will be discarded. If all change in Natural Science was represented on a number line, there are three possible cases:

1. Negative change (change that doesn't refine our understanding of the natural world or produce something of utility)
2. No change
3. Positive change (change that refines our understanding of the natural world or produces something of utility)

Since the scientific community does not value negative change, negative change is ignored. In 2012, Seralini's study that links GMO corn to the development of tumours in rats was published in *Food and Chemical Toxicology* (Séralini *et al.*). However, Seralini's study was determined to be inconclusive since the sample size was too small and the species of rats used were genetically predisposed to developing tumours (Séralini *et al.*). After Seralini's study was proven to be invalid, the *Food and Chemical Toxicology's* editorial board deemed the paper irrelevant, retracting the paper in 2013 (Séralini *et al.*). Thus, as Seralini's study resulted in no actual change within Natural Science, it is discarded.

Moreover, it can be argued that even the smallest positive change is progress. *Nature*, a prestigious scientific journal, constantly publishes new studies that refine our understanding and spark more applications of CRISPR-Cas9, such as how Ma *et al.*'s *MiCas9* increases large size gene knock-in rates and reduces undesirable on-target and off-target indel edits (Ma *et al.*). Since these papers will only have a small reach within the scientific community, they may only be considered as small positive changes. However, *Nature's* editorial board still recognizes their potential worth in progressing the scientific community's understanding and applications of CRISPR-Cas9. Therefore, even the smallest positive change in the Natural Sciences is arguably the same as progress.

In the Arts, progress is a positive difference that, when viewed in retrospect, allows artists to bet-

ter convey the human experience. One way progress may be achieved in the Arts is the creation of new ideas to express the human experience. Duchamp's *Fountain* is a revolutionary artwork where Duchamp signed "R. Mutt, 1917" on a urinal and submitted it to the Society of Independent Artists' salon (Mann). Initially, Duchamp's *Fountain* was rejected by the Society's board as it was not considered a work of art (Mann). However, Duchamp's *Fountain* is now considered "ground-breaking" as he fundamentally questions the definition of Art, suggesting that the artistic message takes precedence over the artwork itself (Mann). In retrospect, it is clear that Duchamp pioneered the artistic expression of a new part of the human experience — reshaping our understanding of what constitutes "Art".

Artistic progress may also be achieved through developing new techniques and media to express the human experience. During the Renaissance, numerous artistic techniques that help make artwork appear realistic were developed, such as foreshortening, sfumato, and chiaroscuro ("Renaissance Painting | Boundless Art History"). Since these techniques allow artists to more realistically express their experiences, the development of artistic techniques qualifies as progress. Artistic progress often occurs in tandem with technological development; an example of this is the development of the artistic medium of film. Following the invention of the first motion picture camera, the first narrative film was Edison's *The Great Train Robbery* (Rosenberg): a 12-minute silent film telling the story of four bandits who robbed a train and were subsequently killed during a shootout (Rosenberg). In retrospect, the creation of Edison's film jump-started the film industry (Rosenberg), providing artists with a new method to express the human experience, such as the crime and death portrayed in his film.

However, since human experience is relative to the audience, what techniques or ideas better convey the human experience in the Arts varies with the audience's geographical location, cultural background, and time period. Released in 2016, Alan Walker's *Alone* is a pop song that explores themes of loneliness (Andrew). Com-

posed in 1880, Tchaikovsky's *1812 Overture* is a piece that commemorates Russia's victory against Napoleon's invasion (Felsenfeld 54). Despite the differences in their messages and musical techniques, which piece is "better" at conveying the human experience differs based on the audience's background. To the Russian audience in 1812, Tchaikovsky's *1812 Overture* may be "better" since not only do the classical techniques such as counterpoint and usage of motifs align with their musical taste (54), but the triumph of winning a war is also a feeling they can relate to. Contrarily, as testified by the more than 1.1 billion views on Youtube (Walker), Alan Walker's *Alone* appeals more to the modern audience since not only is the VI-IV-I-V chord progression simple and catchy, aligning with their musical taste, but loneliness is also a common experience that resonates with a younger audience can resonate with (Cigma). In essence, neither the artistic techniques used nor the aspects of human experience conveyed determine which piece is better, but rather the ability of the piece to resonate with the audience. Thus, due to the relativistic nature of the human experience, it's impossible to judge which song is better at conveying the human experience.

While it is undeniable that knowledge in Natural Science and the Arts has progressed, it is difficult to distinguish between change and progress within these areas of knowledge: In Natural Sciences, all change is a type of progress, since any change that doesn't improve our understanding of the natural world or result in the production of something useful is discarded; in the Arts, all progress is a type of change because what "better" conveys the human experience depends on the audience's geographical location, cultural background, and time period. Mankind devotes a lot of resources to pursue "progress" in areas of knowledge because progress allows us to apply new knowledge to improve our living standards. However, further reflection on knowledge development in different areas of knowledge reveals a fundamental question: should achieving "progress" be the ultimate goal of investing in different areas of knowledge? Should we abandon the Arts because all artistic progress is merely a type of

change? Should we ignore funding branches of Natural Sciences just because they don't result in any direct application? Answering these questions may be the key to minimizing the opportunity cost of the resources we invest in the different areas of knowledge.

Works Cited

- Andrew. "Alan Walker Unleashes Beautiful New Heart Melting Single [Must Listen]." Dec. 2016. <https://www.youredm.com/2016/12/02/alan-walker-releases-new-single-alone/>. Accessed December 16, 2020.
- Bradford, Alina. "Deductive Reasoning vs. Inductive Reasoning." Live Science, July 2017. <https://www.livescience.com/21569-deduction-vs-induction.html>. Accessed December 16, 2020.
- Cigma. "Cigma U.S. Loneliness Index." Cigma, May 2018. https://www.multivu.com/players/English/8294451-cigma-us-loneliness-survey/docs/IndexReport_1524069371598-173525450.pdf.
- Felsenfeld, Daniel. *Tchaikovsky: A Listener's Guide*. Amadeus Press, 2006. 54. Print.
- Ma, Linyuan, *et al.* "MiCas9 increases large size gene knock-in rates and reduces undesirable on-target and off-target indel edits." *Nature Communications* 11 (Nov. 2020). doi:10.1038/s41467-020-19842-2.
- Macdonald, Fiona. "8 Scientific Papers That Were Rejected Before Going on to Win a Nobel Prize." Aug. 2016. <https://www.sciencealert.com/these-8-papers-were-rejected-before-going-on-to-win-the-nobel-prize>. Accessed December 16, 2020.
- Mann, Jon. "How Duchamp's Urinal Changed Art Forever." May 2017. <https://www.artsy.net/article/artsy-editorial-duchamps-urinal-changed-art-forever>. Accessed December 16, 2020.
- "Renaissance Painting | Boundless Art History." lumencandela, <https://courses.lumenlearning.com/boundless-arthistory/chapter/renaissance-painting>. Accessed December 16, 2020.
- Rosenberg, Jennifer. "First Silent Film: The Great Train Robbery." Feb. 2019. <https://www.thoughtco.com/first-silent-movie-the-great-train-robbery-1779195>. Accessed December 16, 2020.
- Séralini, Gilles-Eric *et al.* "Republished study: long-term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize." *Environmental sciences Europe* 26 (2014): 14. Web.
- Vickers, Peter. "The misleading evidence that fooled scientists for decades." Sept. 2018. <https://theconversation.com/the-misleading-evidence-that-fooled-scientists-for-decades-95737>. Accessed December 16, 2020.
- Von Halban, H., F. Joliot, and L. Kowarski. "Number of Neutrons Liberated in the Nuclear Fission of Uranium." *Nature* 143 (Apr. 1939): 680–680. DOI:10.1038/143680A0.
- Walker, Alan. "Alan Walker - Alone - YouTube." Dec. 2016. https://www.youtube.com/watch?v=1-xGerv5FOk&ab_channel=AlanWalker. Accessed December 16, 2020.

Smallpox and its associated political implications in Qing dynasty China

William Tristan Lee III 李欣隆

Introduction

The earliest documented record of smallpox appeared in China in the 4th century CE (Common Era). As a disease, smallpox was so deadly and persistent that it was only finally eradicated in China as recently as the early 1960s (Hui and Yutu 913). Smallpox affected both young and old with symptoms of varying severity. These symptoms included fever, body aches, vomiting and rashes accompanied by pus and scabs. Although a common characteristic among those who recovered from smallpox was an immunity against reinfection (Taub *et al.*, 2008), Qing dynasty China lived under the constant risk of smallpox outbreaks among large populations in Manchuria, the capital Peking, and most significantly within the imperial family. This resulted in serious public health issues, but also had strong political implications (Chang, “Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)” 189). This paper aims to explore the smallpox epidemic during the Qing dynasty by understanding the role it played in various political intrigues, examining the Manchu regime’s implementation of an early quarantine system known as the *bidousuo* (「避痘所」), and its gradual adoption of Chinese inoculation practices. It also considers how China’s early attempts to contain and manage smallpox’s spread through inoculation was perceived by the West.

1 Manchu rule and susceptibility to Smallpox

The Qing dynasty was established by northern Manchu invaders from beyond the Great Wall and was the last Chinese imperial dynasty, reigning from 1644 to 1911. Manchuria was geographically segregated from the rest of China and its natural

landscape, sparse population and nomadic culture allowed Manchus to live a considerable distance from each other. Harsh long winters also made the smallpox virus less prevalent. These conditions meant that Manchus had limited exposure to smallpox compared to the Chinese from the more densely populated south. However, this initial advantage also adversely affected the Manchu community because when smallpox eventually penetrated their communities, the same lack of exposure that had protected the Manchu earlier, effectively made them more susceptible to the smallpox contagion. Later, when the Manchu rode south to conquer China, many more found themselves lacking immunological defenses and facing serious dangers of the disease (“Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)” 179).

With the constant threat of smallpox as a backdrop, the Manchu regime’s primary focus after the 1644 conquest of China was to maintain political supremacy. At the center of the regime’s quest for political supremacy was the importance it placed on safeguarding the imperial family against the ravages of the disease. An example of this importance were the strict rules enforced whenever smallpox outbreaks occurred. This can be evidenced by the *bidousuo* where imperial family members, who had not yet gotten the disease, went to be sequestered during outbreaks. *Bidousuo* effectively served as Qing dynasty “quarantine centers” and were specifically used as places of safety and refuge for non-infected members of the household. As retreat to the *bidousuo* became a regularly enforced practice for protection against smallpox in Manchu-ruled China, an imperial family member’s frequent withdrawals from court could potentially jeopardize

dize his or her social and political status, further inviting the possibility of political posturing, maneuvering and intrigue from rivals within court.

Additionally, as smallpox outbreaks occurred and recurred within the general Manchu population, Manchu authorities, as a means to maintain a firm grip on social and political order, began the common practice of summarily expelling citizens with fever or scabies without verifying whether the illness was smallpox related. To help further contain the spread of the disease, physically weaker smallpox patients were required to move elsewhere to stay outside of the more densely populated capital of Peking. In many cases, poor families who could not afford to leave were forced to abandon their sick children in the streets and some were driven to intentionally kill their children to avoid having to move. As a consequence, the Manchu occupation of Peking also contributed to growing political tensions outside of the Manchu court, as the majority of Chinese inhabitants began feeling alienated by the expulsion regulation and increasingly resented the ethnic minority Manchu rulers for that (“Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)” 190).

Furthermore, at least a century prior to the Manchu Conquest in 1644, the Chinese had already developed highly reliable methods of combating smallpox, yet the Manchus were not necessarily privy to the pre-existing medical knowledge in the early decades of Qing dynasty rule. This lack of medical knowledge, coupled by the political uncertainty, fueled Manchu court intrigue and political instability each time an imperial family or court member became infected with smallpox or came into contact with someone who was sick. It may have also contributed to a sense of apathy towards the Manchu government and increasing unease among the ordinary people, who were left abandoned or banished to live alone in fear and ignorance about the disease (“Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)” 184).

2 Smallpox and Qing Dynasty political intrigue

During the early Qing dynasty, smallpox was thrust onto the political stage and played a vital role within political affairs. During this time, it became increasingly more common for the violation of smallpox regulations (both real and contrived) to be used as a pretext to assert political influence or dominance within the Manchu regime.

As early as 1639, Hong Taiji (Abahai), the founding emperor of the Qing Dynasty, publicly announced that the Manchu prince and military general Dodo had committed a total of seven crimes, one of which was directly related to smallpox. During the conquest of China, Dodo had used segregation in the *bidousuo* as an excuse to avoid going into battle and instead took this time to indulge in a wide range of leisurely activities. While Dodo's six other crimes were eventually forgiven, the one related to smallpox was not. As a result, Dodo was demoted from his positions and was deprived of his property, thus securing Abahai's political authority (“Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)” 190).

Another example of early Qing Dynasty political intrigue that had connections to smallpox involved Abahai's eldest son, Prince Haoge (1609–1648). After Abahai died in 1643, Haoge's intention was to ascend the throne, however his goal was obstructed by Dorgon, Dodo's elder brother. As Abahai's ninth son, Fulin, who would later ascend the throne as the Shunzhi emperor was only six years old at the time, Dorgon solidified his political influence in Shunzhi's court by assuming the title of co-regent. Dorgon's co-regency was shared with Jierhalang, who as the younger brother of Nurhaci, the founder of the Qing dynasty, was a powerful political figure himself. In a politically motivated move to promote his self-image and assert political dominance, Dorgon had accused Haoge of “countless crimes” with those involving smallpox viewed as the most offensive. One of Haoge's many alleged crimes was that Haoge had attempted to undermine Dorgon's bid

for leadership, by invoking the fact that Dorgon had never been infected by smallpox, and that this susceptibility would adversely affect his longevity as a leader. Since the charge was viewed as a significant threat to Dorgon's political power, Dorgon took the opportunity to gain any advantage and was eventually able to impose serious fine on Haoge, in addition to depriving him of his rank and confiscating much of his property ("Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)" 191).

In 1648, soon after Haoge was effectively removed from power and the Shunzhi emperor's reign was firmly established, Dorgon's and Jierhalang's coregency began experiencing periods of political intrigue and upheaval, again, with smallpox being the pretext. In an attempt to curry court favor and assert Dorgon's political dominance, Jierhalang was made to confess to the crime of mishandling the building of a *bidousuo* on a Manchu prince's property, and as a result was stripped of his title. Although this alleged offense was quite serious, it could have also more than likely served as a false justification for punishing Jierhalang for his original allegiance to Haoge, which Dorgon perceived as disloyal and potentially threatening to his own political influence ("[Time Frame and Reasons of Kangxi Emperor Adopted Variolation]" 192).

Dorgon's political maneuverings, particularly from 1644-1650, helped propel him to dominance in the early Qing dynasty. However, for all of Dorgon's efforts, nearly all of his contributions to the Manchu empire were completely discredited after his death in 1650 and his fall from political favor could be tied back to smallpox and Jierhalang, whose demise Dorgon had helped orchestrate. This time, Jierhalang was able to exact his revenge by convincing Emperor Shunzhi that Dorgon had misappropriated funds to build a *bidousuo* for him in Nanyuan and in doing so had also reduced the number of imperial bodyguards there. By doing this, Dorgon was accused of having both overstepped imperial directives and compromised the emperor's safety. Although prior to his death, while Dorgon was still held in high political

esteem, building the *bidousuo* had been viewed as a gesture to honor and protect the emperor, after Dorgon's death these good intentions were viewed suspiciously as an act of sabotage rather than devotion to the Shunzhi emperor. Therefore, upon Dorgon's death, instead of being celebrated as a faithful servant, he was not only accused of exceeding his authority by sanctioning the Nanyuan *bidousuo*, but also exposed as harboring evil intentions to harm the emperor when they traveled away from the city to isolate ("Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)" 192).

Political intrigues connected to smallpox continued occurring within the Manchu court towards the end of the 18th century. Shortly after Emperor Qianlong's death in 1799, Heshen, the highest-ranking court official was issued a death sentence and accused of 20 crimes, at least one of which was for an offense related to smallpox. The offense alleged that Heshen had altered the new emperor Jiaqing's imperial order of allowing high ranking Mongolian official who had not yet had smallpox not to attend Emperor Qianlong's funeral. Instead, as the allegation went, Heshen ordered all high-ranking official to come, regardless of whether they had previously been infected by the disease, thereby potentially compromising the new emperor's health and destabilizing the new imperial regime. While Jiaqing may have just seized an opportunity to eliminate one of the last vestiges of the Qianlong reign, Heshen was charged with ignoring diplomatic policy and intending to harm the country. As a consequence, Heshen's powers were removed, and his property was taken over by the emperor ("Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)" 192).

3 Self-protective measures and growing acceptance of Inoculation

By the early Qing dynasty, concerns about smallpox and its deadly outcomes had already affected

thousands of people in Manchuria and Peking, which helped pave the way for conditions that would eventually support a growing willingness to adopt inoculation practices. In fact, smallpox was the first pandemic where people applied the inoculation practice as a mechanism to combat disease by intentionally exposing themselves (National Museum of American History). Sinologist Joseph Needham argued that the first evidence of inoculation for smallpox being practiced in China dated back to the 10th century (Needham and Sivin 154). However, more credible sources indicate that smallpox inoculation was not as scientifically well-developed until hundreds of years later until the early 16th century (Chang, “[Time Frame and Reasons of Kangxi Emperor Adopted Variolation]” 30-32). This inoculation process required healthy people to inhale powder made from crusted smallpox scabs or by inserting pus from a smallpox pustule into a flesh wound or scratch made by a sharp instrument (See Fig 1). This method led to people to develop mild symptoms of the infection, even though some could develop much more severe symptoms and even occasionally, result in fatal cases. However, those who did recover were usually immune to any subsequent smallpox exposure (Science Museum). Of those who had been inoculated, the mortality rate was around 1% to 2%, which was far better than the rate of unchecked, naturally occurring smallpox, which was 20%-30% (Needham and Sivin 149).

By the end of the 17th century, Manchu rulers’ acute fears about smallpox and their interest in self-protection had, supported a gradual acceptance of inoculation. A major turning point came in 1661 and the years immediately following a series of historically significant events. The first was Emperor Shunzhi’s death from smallpox at the age of 22, despite his best efforts to avoid it. The second event came when Shunzhi’s successor and second eldest son, Emperor Kangxi, was chosen to ascend the throne largely due to the fact Kangxi had already survived childhood smallpox, making him immune to the disease for life. The last event came as Emperor Kangxi began mandating and publicly endorsing compulsory smallpox inoculation for his family and all those resid-



Fig. 1. Illustration of Inoculation Practice in the Qing Dynasty. Koon, Wee Kek. Powdered Pus up the Nose and Other Chinese Precursors to Vaccinations, SCMP, 6 Apr. 2020, www.scmp.com/magazines/post-magazine/short-reads/article/3078436/powdered-pus-nose-and-other-chinese-precursors.

ing within the Imperial Palace. Although there was certainly some immediate risk of acute infection, the potential advantage of lifelong protection against smallpox was an attractive prospect (Needham and Sivin 149).

It was also during this time when Kangxi’s interest in inoculation coupled with the Manchus’ and Mongols’ intense fear of smallpox intersected to help create an alternative means of maintaining the smooth and uninterrupted conduct of official court business. Normally, court decree mandated that visiting officials were to be received at court. However, this meant that Manchu and Mongol officials who were mostly situated in China’s far north, were required to travel a considerable distance to the capital. This greatly heightened the risk of contracting smallpox for the visiting dignitary, his entourage, and the imperial court. To safeguard against a smallpox outbreak within court, while also promoting the ease of conducting diplomatic affairs between the regions, Emperor Kangxi continued to receive *shoushen* 「熟身」 leaders (those who had already had smallpox) at court, but hosted *shengshen* 「生身」 (those who had not yet had smallpox) at facilities he had built close to the northern border of the Great Wall, which included a large hunting park, Mulan Weichang 「木蘭圍場」 and a summer villa. This alternative al-

lowed court visitors to bypass the capital entirely and significantly reduced all parties' chances for smallpox exposure. It also provided the Manchus the unencumbered means to carry out a myriad of political affairs within court without the imminent threat of a smallpox outbreak, further securing the Manchu minority's political stronghold (Chang, "[Time Frame and Reasons of Kangxi Emperor Adopted Variolation]" 30-32).

Although inoculation had already been widely practiced and promoted in China since well over a hundred years earlier, English physician and scientist, Edward Jenner was ultimately credited for developing a less dangerous form of immunization against smallpox in 1796, the vaccination. China's inoculation practice and the Jennerian vaccination were similarly effective, the main difference being that Chinese inoculation involved using the pus or the scabs from smallpox patients, while Jenner's vaccination used matter from the milder cowpox virus (Housman). Although the Chinese inoculation technique of inhaling powder made from the crusts of smallpox scabs had reached England in the 1700s, Jenner's method eventually became the globally accepted practice (Boylston 309-313).

4 Western misconceptions of Smallpox in China

Qing dynasty leaders undertook significant efforts to cope with smallpox, and at times, also used smallpox as a means for political leverage. Also noteworthy during this period were some of the West's diverging philosophical views about China, some of which were rooted in the West's misconceptions about China's inoculation methods during a time when much of the world was still struggling to understand smallpox.

An example of this diverging view was evidenced through Pierre Martial Cibot (1727-1780), a well-known Jesuit missionary, who lived at court in Qing dynasty China from 1760-1780. By the time Cibot left France for Peking in 1758, inoculation had already become the subject of heated debate between the French church and Enlightenment thinkers. This debate centered on how inoculation

involved deliberately infecting people with the disease to stimulate an immune response. There were those who believed the method to be unethical because inoculation could sometimes cause death, while others were quick to refute this position and argued for inoculation as an acceptable means to protect against the smallpox spread. Cibot argued that the continued presence of smallpox in China alone proved that inoculation had failed, and that China was indeed the land of disease and not cure. He also went as far to refer to Chinese method of inoculation as "pathetic stupidity" and "lunacy and inconsistency" (Heinrich).

While some of the methods used to manage smallpox, particularly during the early Qing dynasty were not necessarily viewed as compassionate or even humane, they were in fact, more varied, uniformly applied and tested than anywhere in the western world. Furthermore, although the Manchus during the early Qing dynasty may not have been as familiar with specific smallpox inoculation methods used by their Chinese counterparts, as early as 1622, more than two decades before the Qing Dynasty began in 1644, Manchus had already implemented tracing systems to report anyone exhibiting smallpox symptoms. Safety guidelines were also established to prevent the spread of smallpox when offers of tribute were brought from visiting dignitaries and when arranging audiences with the emperor (Heinrich). Therefore, the popular western notion perpetuated by Cibot was at best incomplete, as it did not fully appreciate early Qing dynasty inoculation practices or its effectiveness, nor did it make reference to the Manchus systematic tracing method which helped contain the spread of smallpox within their communities.

Cibot's accounts, which by the 19th century had been widely circulated in Europe, helped contribute to growing views in the West of China as the "Sick Man of Asia." This term was used both literally to describe poor health, but also figuratively to suggest poor governance. Although Cibot was considered a western "authority" on China at the time, the term fed into an increasingly fragmented narrative, which led many Europeans and Americans into believing that China was the "cra-

lent with the people of Europe, as among those of the Centre Kingdom” (Heinrich).

Conclusion

Although eradicated since the early 1960s, smallpox is still recognized as one of the deadliest diseases in global history. In Qing dynasty China, smallpox posed a significant, if not deadly, threat to the Manchus, which made it a convenient pretext for sparking political intrigues and ongoing contentiousness at court involving emperors, high level military officials and dignitaries both centrally and peripherally. In time, Qing dynasty emperors implemented strict quarantine-like measures such as the *bidousuo* and adopted the Chinese practice of inoculation, which supported improved public health across the empire. While limited and somewhat dubious accounts in the West concerning the efficacy of these preventative practices helped perpetuate certain misconceptions and contributed to negative perceptions of China, it is without question that countless lives were spared over 150 years before Edward Jenner’s vaccine arrived on the scene to achieve similar effects around the world.

Works Cited

- Boylston, Arthur. “The origins of inoculation”. *Journal of the Royal Society of Medicine* 105 (July 2012): 309–313.
- Chang, Chia-Feng. “Disease and Its Impact on Politics, Diplomacy, and the Military: The Case of Smallpox and the Manchus (1613-1795)”. *Journal of the History of Medicine and Allied Sciences* 57 (Apr. 2002): 177–197.
- Chang, Chia-Feng. “[Time Frame and Reasons of Kangxi Emperor Adopted Variolation]”. *Zhonghua Yi Shi Za Zhi (Beijing, China : 1980)* 26 (1996): 30–32.
- Heinrich, Ari Larissa. “Before coronavirus, China was falsely blamed for spreading smallpox. Racism played a role then, too.” May 2020. <https://theconversation.com/before-coronavirus-china-was-falsely-blamed-for-spreading-smallpox-racism-played-a-role-then-too-137884?> Accessed 17 November, 2021.
- Housman, Talya. “Variolation vs. Vaccination: 18th Century Developments in Smallpox Inoculation | Beehive”. May 2020. <https://www.masshist.org/beeiveblog/2020/05/variolation-vs-vaccination-18th-century-developments-in-smallpox-inoculation/>. Accessed 17 November, 2021.
- Hui, Xu and Jiang Yutu. “The eradication of smallpox in Shanghai, China, October 1950—July 1951”. *Bulletin of the World Health Organization* 59 (1981): 913–917.
- National Museum of American History. “History of Vaccines”. Feb. 2005. <https://amhistory.si.edu/polio/virusvaccine/history.htm>. Accessed 17 November, 2021.
- Needham, Joseph and Nathan Sivin. “The Origins of Immunology”. Cambridge University Press, 2000. Science and Civilisation in China. Volume 6, Biology and Biological Technology. Part VI, Medicine.
- Science Museum. “Smallpox and the story of vaccination”. Apr. 2019. <https://www.sciencemuseum.org.uk/objects-and-stories/medicine/smallpox-and-story-vaccination>. Accessed 17 November, 2021.
- Taub, Dennis D., *et al.* “Immunity from Smallpox Vaccine Persists for Decades: A Longitudinal Study”. *The American Journal of Medicine* 121 (Dec. 2008): 1058–1064.

A data-driven study on the effects of changes in local sea surface temperature on the health of coral reefs in Hoi Ha Wan, Hong Kong

Julie Tam 譚幸臨

1 Introduction

1.1 Background Knowledge

Scleractinia (phylum Cnidaria) are reef-building corals found in tropical & subtropical oceans, composed of sedentary polyps. Heterotrophic coral polyps, besides being filter feeders, form a mutualistic endosymbiotic relationship with photosynthetic dinoflagellates zooxanthellae such as *Symbiodinium microdriaticum* (Vernon *et al.*, 2016). Zooxanthellae are the primary nutrient source, supporting coral survival in oligotrophic seas.

Increased sea temperatures induce stress in corals, bleaching is a stress response that results when the coral-algae relationship breaks down. Zooxanthellae are unable to process incoming light without releasing harmful oxygen radicals, so will either degenerate in the tissue or are released from the coral polyps (Marshall *et al.*, 2006). The loss of zooxanthellae makes the coral tissue appear unpigmented with the white CaCO_3 skeleton being clearly visible, a phenomenon known as coral “bleaching”. Bleached corals are still living and if stressful conditions subside, zooxanthellae can repopulate their tissues and the corals can survive the bleaching event. The inability to recover from the loss of zooxanthellae leads to coral mortality.

Hong Kong is located on the southeast coast of China, to the east of the Pearl River estuary (see Figure 1), bordered by the South China Sea. The salinity and temperature of Hong Kong seawater are affected by multiple factors. Hong Kong’s subtropical climate brings hot, rainy summers with monsoons, which contrast with cool, dry winters (Cope, 1986).

During Hong Kong’s summer monsoon season, the warmer Hainan Current flowing from the south is

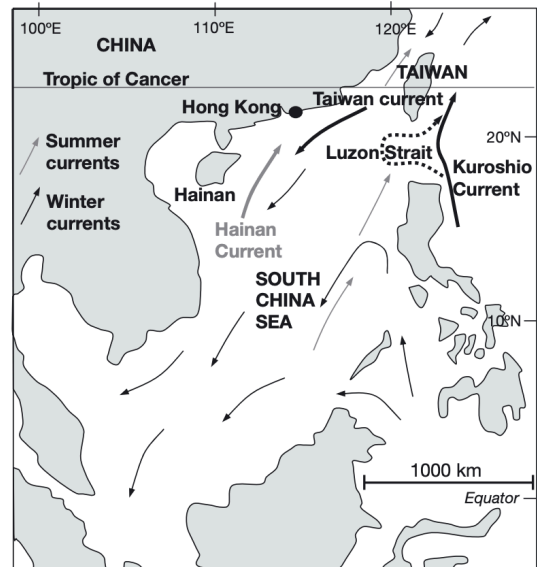


Fig. 1. Map showing the two ocean currents affecting Hong Kong – the Hainan current and the Kuroshio current (Goodkin *et al.*, 2011).

the predominant ocean current. Whereas during the winter monsoon, a cold current – the Kuroshio Current as well as the Taiwan current flowing from the north – influence Hong Kong’s sea temperature. The lowest sea temperature during winter is around 15 °C, while the highest sea temperature can reach 30 °C and higher. Hong Kong coral reefs accommodate a wider sea surface temperature (SST) range than most reefs in other locations (Bradbeer *et al.*, 2008).

Hoi Ha Wan was one of the first Marine Parks established in Hong Kong, founded in 1996. It is a sheltered bay located north of the Sai Kung West Country Park; 64 out of the 84 stony coral species recorded in Hong Kong can be found here.

1.2 Significance and worthiness of investigation

Coral reefs only make up less than 0.2% of the world’s oceans but provide habitat to over a quar-

ter of all marine species (Thornhill, 2012). The scarcity of corals is due to their specific habitat requirements, including temperature, salinity, nutrient availability, turbidity, and light availability (Guan *et al.*, 2015).

The complex structures of hard corals support a diverse range of marine organisms, creating biodiverse ecosystems that provide goods and services on which more than 500 million people rely, as a source of food, for income and coastal protection (Orlowski & Rhodes, 2017; Hoegh-Guldberg, Pendleton, & Kaup, 2019). The fifth Global Environment Outlook (GEO 5) predicts that if global and local threats to corals are left unchecked, it could result in mass coral mortality by 2050 (United Nations *et al.*, 2012; Tamelander, 2016). The loss of coral cover would result in shifts from coral-dominated to algal-dominated reefs, consequently leading to the loss of the goods and services (Tamelander, 2016).

The 2013 Intergovernmental Panel on Climate Change (IPCC) report predicts a 3–4 °C increase in the global SST by 2100 if current greenhouse gas patterns are left unchanged (Environmental Protection Department, 2020). Marshall *et al.* (2006) suggested that the primary cause of mass coral bleaching events is due to increased sea temperatures. Sea temperature increases of 1–2 °C above the long-term average maximum can trigger mass coral bleaching (Marshall *et al.*, 2006). This information about how coral bleaching is caused by an increase in sea temperature led me to believe that all coral reefs are extremely susceptible to an increase in temperature. I wanted to investigate whether this was occurring in the marginal reefs of Hoi Ha Wan, Hong Kong.

Goodkin *et al.* (2011) compared SST ranges of marginal reefs found globally, reefs in Hong Kong and the Persian Gulf had the highest SST range. Similarly, Bradbeer *et al.* (2008) has found that due to Hong Kong's geographical location, hydrography and sub-tropical climate, Hong Kong's corals are more stress-resistant than corals in tropical climates. As Marshall suggested, corals and coral reefs have survived a variety of changes in their physical and chemical environment over the past half-million years, and they are unlikely to disap-

pear altogether, even under extreme climate variations. However, reef health is likely to significantly deteriorate as a result of coral bleaching (Marshall *et al.*, 2006).

Thompson (2019) conducted *in vitro* studies on coral metabolism mechanisms to investigate the physiological responses of corals to seasonal changes (including temperature), with the most commonly found species of corals found in Hong Kong. Corals from Hong Kong are an excellent model for studying how corals cope with changes in climate, as corals acclimatized to seasonal variation could extend or compress their thermal ranges. Despite the extreme thermal variations, Hong Kong is home to a relatively healthy coral community of more than 90 species (Thompson, 2019).

I wanted to investigate whether changes in local SST would correlate with the health of two coral reefs in the government designated Marine Park, Hoi Ha Wan Marine Park in Hong Kong.

2 Investigation

2.1 Hypothesis

H_i : The effect of increases in Hong Kong's annual SST above the optimum temperature range will cause a decrease in coral health from two coral reefs in Hoi Ha Wan, Hong Kong, showing an inverse linear relationship. This will support findings by Marshall *et al.* (2006), suggesting that increased SST breaks down the mutualistic relationship between coral polyps and zooxanthellae, inducing coral bleaching as a stress response. When sea temperatures are too high and exceed optimal temperature ranges, enzymes involved with the Krebs Cycle (associated with cellular respiration) denature, which restricts energy transformation (Thompson, 2019). Zooxanthellae will be expelled, coral polyps will then rely on heterotrophic feeding and stored lipids (Jokiel, 2004).

H_0 : Sea temperature does not affect coral health. This is because the annual SST increases above the corals' optimum thermal ranges do not induce thermal stress in corals found in Hoi Ha Wan, Hong Kong. Zooxanthellae can maximize their capacity

for photosynthesis via acclimation depending on the light and temperature regimes of their environments (Jokiel, 2004).

2.2 Independent Variable

The independent variable of this investigation is local monthly SST, measured in °C, collected through a secondary source. There are two data sets available for SST data in Hong Kong.

The first data set is from the Hong Kong Observatory (HKO) Monthly Data for Single Element web page (Hong Kong Observatory, 2019). Sea temperature is collected by HKO at their North Point station in the morning. Sea temperature is collected daily and a monthly average is calculated, data is available from August, 1974 until July, 2020.

The second data set is from the Hong Kong Environmental Protection Department (EPD), their Marine Water Quality Data web page provides sea temperature data specific to 76 different marine water monitoring stations in Hong Kong (as shown in Figure 2). The monitoring station closest to Hoi Ha Wan is station MM6. Water is sampled once per month using a rosette sampling system. Water samples are taken at three depths– “surface” water samples are collected one metre below the surface of the water; “middle” samples are collected midway between the surface & the sea bed; “bottom” samples are collected one metre above the sea bed (Wong, 2006). Data for water samples taken at the “surface” depth was selected as coral reefs in Hoi Ha Wan are shallow, around 2-5 metres deep. Water samples at the “middle” depth would be around 8-10 metres deep, as Hong Kong’s Eastern waters “average between 20 and 30 metres in depth” according to the EPD (Wong, 2006). The use of a Differential Global Positioning System (DGPS) ensures the location accuracy of ± 1 metre for water samples (Wong, 2006). Monthly SST data is available from 1986-2018.

The HKO data set provides monthly average SST values, extreme and drastic temperature values that can change within days are smoothed out by the monthly average (Guan, Hohn, & Merico, 2015). While the EPD data set just collects water samples once per month, and on different days in each month. Collecting SST data once per month

does not record the extreme and drastic temperature fluctuations that can occur within days. The EPD also lacks data for January, 2019 to July, 2020.

Among the two data sets, the EPD data set is more appropriate for this investigation as the water sampling station is specific to the location of coral reefs in Hoi Ha Wan, as compared to the HKO data set where the water is sampled in North Point (shown in Figure 5). The location specificity of the SST data makes the EPD data set more relevant. North Point is located in Victoria Harbor which is not exposed to the ocean currents affecting the Eastern waters of Hong Kong.

To use the EPD data set, the data must be extrapolated to create a representation of SST trends from January, 2019 to July, 2020. Firstly, the EPD data set is compared to the HKO data set to identify congruence. Both data sets show congruence in the yearly SST fluctuations from January, 2015 to December, 2018 (shown in Figure 3). The seasonal temperature cycle for both data sets oscillates annually in similar patterns. It may be predicted that the monthly temperature in the missing period, from January 2019 to September 2020 of the EPD data set, will follow the pattern of the previous four years.

Next, the data of the same month in all available years, (for example, SST from the EPD data set for October in 2014, 2015, 2016, 2017, 2018, were added up and averaged into a five-year average October SST value). The same approach was applied to the HKO data set. The five-year average SST value for each month in both data sets was graphed together as in Figure 4.

Closely congruent SST fluctuations due to Hong Kong’s seasonal changes in climate for Hoi Ha Wan and North Point can be identified. Both data sets have an annual SST minimum in February (minimum SST of 15.9 °C), an increase from February to July, a maximum SST in July & September (maximum SST of 27.9 °C), a decrease in SST in August, then a decrease in SST from September to December. Therefore, the five-year average SST value for each month can be used to extrapolate the missing two years (January, 2019 to July, 2020) in the EPD data set as shown in the graph below (Figure 4), with the red line representing the extrap-

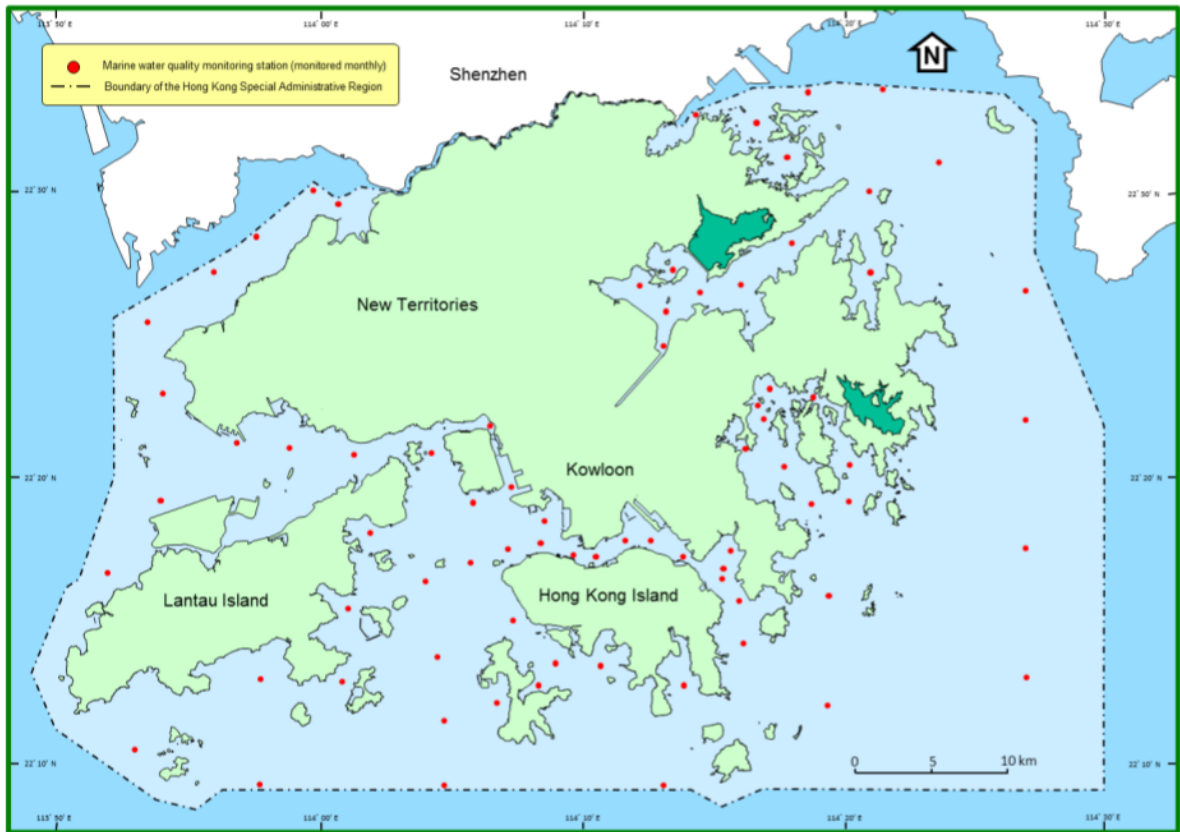


Fig. 2. Map of EPD marine water quality monitoring stations in Hong Kong (EPD).

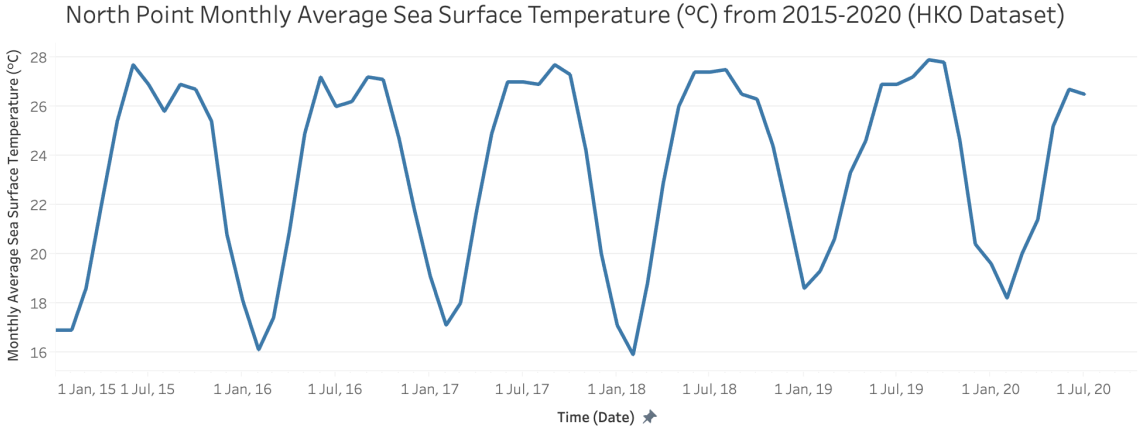


Fig. 3. Graph comparing the monthly SST data (°C) from the EPD data set (from 2015-2018) to the HKO data set (2015 to 2020).

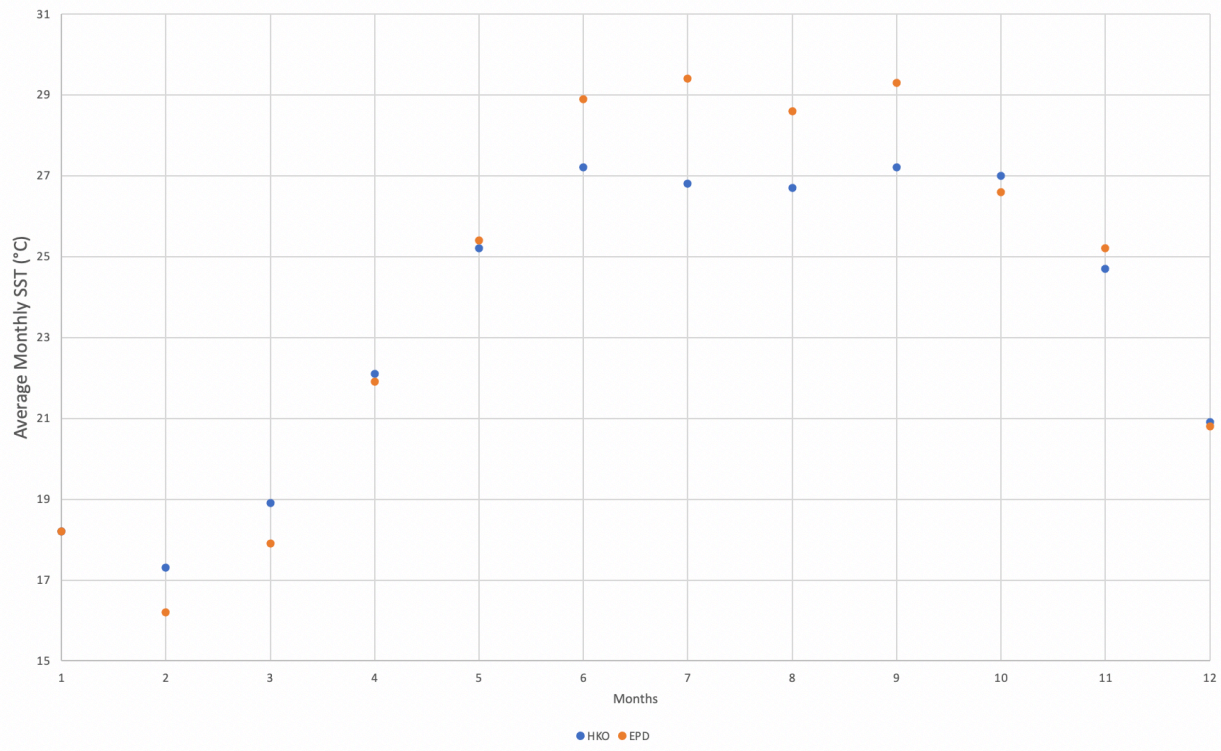


Fig. 4. Graph showing the five-year average monthly SST (°C) for the EPD and HKO data sets.

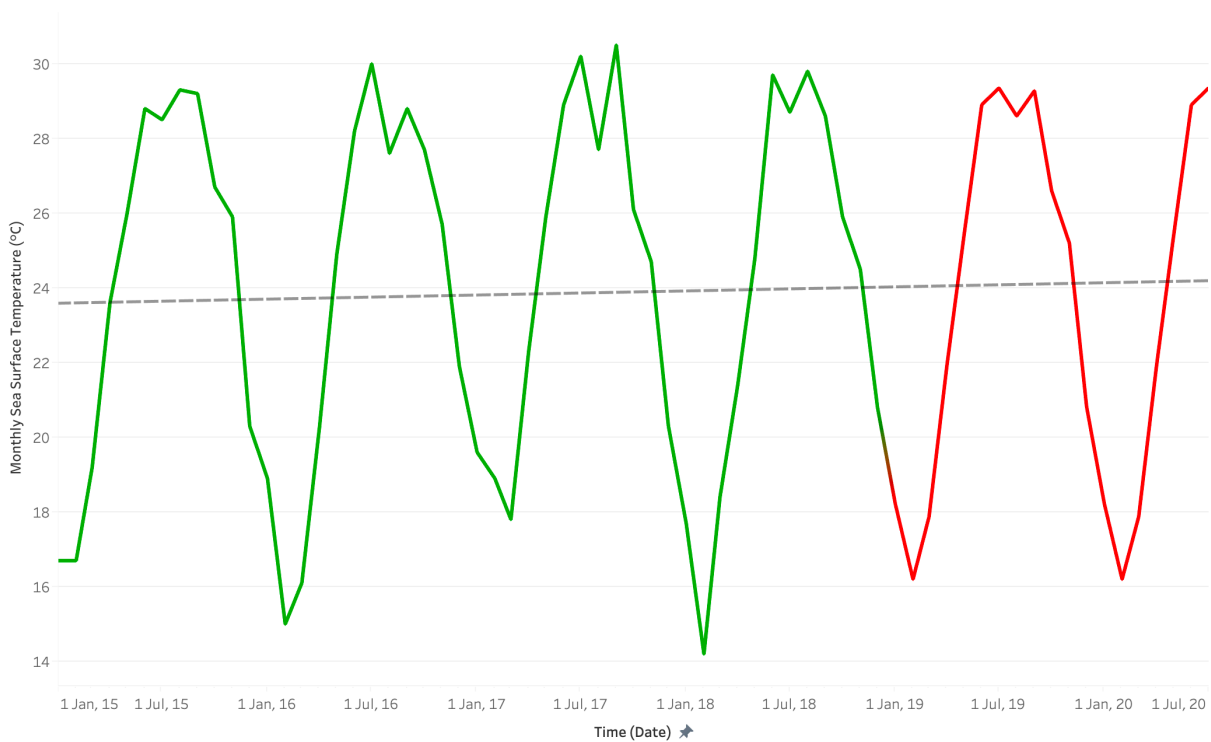


Fig. 5. Hoi Ha Wan monthly SST (°C) from January, 2015 to July, 2020 with extrapolation from the EPD data set.

olated data. However, using existing SST data to project subsequent yearly SST fluctuations cannot reflect any unconventional deviations or anomalies of SST changes that have happened in 2019-2020.

2.3 Dependent variable

The dependent variable of this investigation is coral health, quantified using a cumulative average (CA) value from the CoralWatch coral health chart. Coral health is measured using the Traditional Chinese CoralWatch coral health chart (specifically created for citizen science programs

in Hong Kong coral reefs, the colors shades slightly differ from the general chart – see Figure 6) created by the University of Queensland. The Coral-Watch coral health chart is a method of quantifying zooxanthellae abundance as an indicator of coral health, not a direct measure of coral health based on coral color.



Fig. 6. The traditional Chinese CoralWatch coral health chart (Chan Hodgson, 2018).

The data collection method uses a line transect to sample 20 corals, the lightest and darkest color of each coral is recorded, corresponding to a number. According to the CoralWatch site, “The chart is a proxy indicator of symbiont density with a 6-point numerical scale. The brightness corresponds to the symbiont density and the four common coral colors assist selecting the accurate number”. The greater the number, the brighter the color of the coral, indicating a greater abundance and density of zooxanthellae, inferring a healthy symbiotic relationship between zooxanthellae and coral polyps (Chan Hodgson, 2018). An average of the darkest & lightest color for each coral is calculated, and then a cumulative average of all the average CA values is calculated for each month of each year. The CA value has a maximum of 6, inferring minimum environmental stress, and a minimum of 1, inferring maximum environmental stress. The method of data collection is based on the reef check method used by Oceanway Corporation Limited, based on the Hong Kong Agriculture,

Fisheries and Conservation Department (AFCD) reef check method.

Coral health data is collected mainly through a secondary source. Historical Coralwatch data from the protected areas, Coral Beach and the Pier Area in Hoi Ha Wan, dating back to 2015 has been provided by Paul Hodgson from Oceanway. Primary research was also able to contribute to one data point for coral health in July, 2020. The main use of fieldwork for primary research is to understand the methodology of how the secondary data was collected for the historical data.

2.4 Controlled Variables

The dive locations chosen for data collection had the same GPS locations as dives conducted by Oceanway Corporation Limited for historical data collection. The exact coordinates of Coral Beach are 22°28'14.8"N, 114°19'54.3"E, and the coordinates for the Pier Area are 22°27'48.3"N, 114°19'50.2"E. The dives were conducted on the same day, 30 minutes apart so the weather, air and sea temperature, visibility conditions recorded were similar. The diving depth was controlled and ranged from 3-5 metres, which was easy to maintain as the depth of coral reefs in Hoi Ha dive sites are a constant depth of around 4 metres.

The coral sample size was controlled: 20 corals per dive site were sampled for data collection, each coral 1 metre apart on a 20-metre transect line. The sampling method was controlled by using a 20-metre line transect laid out underwater, two divers sampled the nearest coral every metre, but one diver starting at 0.5 m and the other diver starting at 0.0 m. This ensures the same coral will not be sampled more than once, and the survey can be employed under turbid field conditions (Cope & Morton, 1988).

The sampling technique was controlled by using the traditional Chinese CoralWatch chart for sampling. The darkest and lightest shade and color of coral were recorded using the corresponding letters and numbers shown on the slate. The tips of the corals were not sampled, coral tips are naturally lighter as it is the area of coral growth, so they are not a representation of coral bleaching/blanching. Visibility was poor due to a high

sediment concentration, a torch was used when the color of coral was hard to determine, due to the poor visibility. Sunlight was not consistent as the weather varies on different data collection days (for historical data). During cloudy periods, a torch was used to accurately determine the coral color.

2.5 Cofounding Variables

Confounding variables include natural and anthropogenic processes. Coral bioerosion (breakdown of coral substrates) was a natural process that occurred due to overgrazing from corallivorous species, including *Diadema setosum* (long-spined sea urchin) and *Drupella rugosa* (gastropods). The lack of *Choerodon schoenleinii* (black-spot tuskfish)– a natural predator for *Diadema setosum*, most likely due to overfishing (Cornish, 2003), has greatly increased *Diadema setosum* populations and abundance (Lam, Shin, & Hodgson, 2007).

A high abundance of sea urchins has a negative effect on the health of corals. The distribution and density of the corallivorous sea urchin *Diadema setosum* has increased in Hoi Ha Wan, causing serious bioerosion in approximately 300 colonies in 2006 (Lam et al., 2007).

In 2006, bioerosion by these 2 species in Coral Beach was so severe that the Agriculture, Fisheries and Conservation Department of Hong Kong (AFCD) removed around 20,000 each of *Drupella rugosa* and *Diadema setosum*, and the area was closed off to human activities until 2007 to promote coral growth. Based on qualitative observations, Coral Beach has recovered from the bioerosion, however, *Diadema setosum* populations are still high.

Another natural process includes temperature changes caused by the El Niño Southern Oscillation (ENSO) cycle, a quasi-periodic cycle of warming and cooling in the tropical central and eastern Pacific Ocean, which impacts weather, sea temperatures and climates worldwide. El Niño is the negative or the warm phase of the El Niño Southern Oscillation (ENSO) cycle, and La Niña is the positive, or the cool phase (Holland, 2009).

El Niño is a period of above-average SST, whereas La Niña is a period of below-average SST, with an anomaly of ± 0.5 °C, according to the National Oceanic and Atmospheric Administration (NOAA) Oceanic Niño Index (ONI). Based on the temperature anomalies from the NOAA ONI, the most recent El Niño occurrences were from November, 2014 to May, 2016, from October, 2018 to June, 2019, and from November, 2019 to March, 2020. The most recent La Niña occurrences were from August, 2016 to December, 2016, and from October, 2017 to March, 2018.

Increased sea temperatures associated with El Niño have caused coral bleaching events, the first widespread mass bleaching occurred during the 1982-83 El Niño, and global mass bleaching was also evident in the 2014-16 El Niño (Marshall, Schuttenberg, West, & Great Barrier Reef Marine Park Authority, 2006).

Anthropogenic processes cause physical disturbances. The use of trawling nets, boat anchoring and blast fishing cause direct physical damage to coral, these disturbances also stir up sediments which have a negative impact on coral health. According to Schmidt (2008), “Bleached reefs, weakened from zooxanthellae loss, are especially vulnerable to the stress of an increased sediment load. Corals are spending their remaining limited energy on cleaning out sediments instead of capturing food. Sediments coat the skeletal reef surfaces that coral larvae would usually settle on and populate”. An increased sediment load had negative effects on already bleached reefs.

3 Method

3.1 Experimental method

Extrapolation of SST data was conducted to create a trendline for monthly SST trends from January, 2015 to July, 2020 for the EPD data set, as outlined in Figure 5.

On the day fieldwork was conducted, dive equipment was first collected and brought onto the dive boat. Second, the conditions for weather, air temperature, sea temperature and visibility of the water were collected on a slate. Then, the GPS loca-



Fig. 7. Map of Hoi Ha Wan Marine Park with the coordinates of the two dive locations where the transects were taken (AFCD Hong Kong).

tion of the dive site was collected via Google Maps, chosen based on the location where Oceanway collected coral health data for their data set.

At the first dive site, Coral Beach (see Figure 7), a 20-metre transect line was carefully laid out underwater, to ensure it did not entangle and damage any coral. 20 corals were examined *in situ* with the assistance of two divers and recorded the coral color every one metre, with one diver starting at the 0m mark and the second diver starting at 0.5 m. The color found on each coral was matched with the square with the most similar color and brightness on the traditional Chinese CoralWatch coral health chart (see Figure 6). This color was recorded on CoralWatch data sheets (see Figure 8). Care was taken to avoid measuring the color of the tips of corals, since there can be a delay in coloration due to the required up-take time for zooxanthellae to enter coral tissue (Lam *et al.*, 2007). After a 30-minute surface interval for decompression time, the same data collection process was repeated at the second dive site, the Pier Area (shown in Figure 5).

4 Results

4.1 Processed Data

Extrapolating average SST value for each month

By extrapolation, the average SST value for each month is given by equation 1.

$$\frac{\sum(SST \text{ values of one month from 2015 to 2020})}{\text{Number of SST values added}} \quad (1)$$

A sample calculation for the average SST value for January (EDP data) is shown in equation 2.

$$\frac{16.7 + 18.9 + 19.6 + 17.7}{4} = 18.2 \quad (2)$$

Average value of coral health of an individual coral

The average value of coral health of an individual coral is given by equation 3.

CORALWATCH DATA SHEET

Group name: _____ Your name: _____
 Email address: _____
 Participation field: dive centre / scientist / environmental / school or university / tourist
 Country of reef: _____ Reef name: _____
 GPS if possible: _____ Sea temperature: _____ °C
 Date of survey: _____ / _____ / _____ Time collected: (ie.14:00 or 2pm) _____
 Weather: sunny / cloudy / raining Your activity: reef walking / snorkelling / diving

***Please note: data will not be accepted on the website if any of these fields are left blank**

Coral Number	Colour Code		Coral Type				Check out these resources...
	L=Lightest	D=Darkest	Br=Branching	Pl=Plate	Bo=Boulder	So=Soft	
example	L: D2	D: E5	Br	Bo	Pl	So	Reid, C., Marshall, J., Logan, D., Kleine, D. (2009) Coral Reefs and Climate Change: the guide for education and awareness. CoralWatch, Brisbane. Siebeck, U.E., Marshall, N.J., Kluter, A. and Hoegh-Guldberg, O. (2006) <i>Coral Reefs</i> 25(3):453-460
1	L: D:	D: D:	Br	Bo	Pl	So	
2	L: D:	D: D:	Br	Bo	Pl	So	
3	L: D:	D: D:	Br	Bo	Pl	So	
4	L: D:	D: D:	Br	Bo	Pl	So	
5	L: D:	D: D:	Br	Bo	Pl	So	
6	L: D:	D: D:	Br	Bo	Pl	So	
7	L: D:	D: D:	Br	Bo	Pl	So	
8	L: D:	D: D:	Br	Bo	Pl	So	
9	L: D:	D: D:	Br	Bo	Pl	So	
10	L: D:	D: D:	Br	Bo	Pl	So	
11	L: D:	D: D:	Br	Bo	Pl	So	
12	L: D:	D: D:	Br	Bo	Pl	So	
13	L: D:	D: D:	Br	Bo	Pl	So	
14	L: D:	D: D:	Br	Bo	Pl	So	
15	L: D:	D: D:	Br	Bo	Pl	So	
16	L: D:	D: D:	Br	Bo	Pl	So	
17	L: D:	D: D:	Br	Bo	Pl	So	
18	L: D:	D: D:	Br	Bo	Pl	So	
19	L: D:	D: D:	Br	Bo	Pl	So	
20	L: D:	D: D:	Br	Bo	Pl	So	

Fig. 8. CoralWatch data collection sheet used for data collection (CoralWatch).

$$\frac{(\text{Number of the darkest coral color} + \text{Number of the lightest coral color}) \cdot \frac{1}{2}}{2} \quad (3)$$

A sample calculation for the average value of coral health of coral no. 1 from Coral Beach, July 2020 is shown in equation 4.

$$\frac{5 + 2}{2} = 3.5 \quad (4)$$

Cumulative average of a coral's health value

The cumulative average of a coral's health value is given by equation 5.

$$\frac{\sum(\text{Average coral health of all 20 corals})}{20} \quad (5)$$

A sample calculation for the cumulative average of coral health for Coral Beach, July 2020, where $\sum(\text{Average coral health of all 20 corals}) = 60.5$, is shown in equation 6.

$$\frac{60.5}{20} = 3.025 \quad (6)$$

r^2 Value

r^2 value for a linear regression line was generated using the software *Tableau*.

Monthly coral health data fluctuates with corresponding monthly SST data in a cyclical pattern each year, as seen in Figures 9 and 10. The lowest coral health CA value (larger CA indicating higher zooxanthellae abundance, inferring greater coral health) can be found during January to February each year, also the months with lowest SST.

In Figure 9, coral health increases and decreases synchronously with the yearly SST fluctuations in 2015 to 2016, where the peak in coral health matches with the peak in SST (in July, 2015 and September, 2016). Starting in 2017 until the end of 2018, there is a delayed response in coral health to temperature fluctuations, where the peaks in coral health do not match the peak in SST. An outlier is the lowest coral health value within the 5-year period, 2.925 in April, 2020, which does not match with the increasing SST trend.

In Figure 10, there is a large and sudden drop in coral health in July, 2015, which decreased to a CA value of 3.4. The lowest and highest CA values observed within the 5-year period are 2.55 in February 2016 and 4.55 in September 2019 respectively. Starting in 2016 until the end of 2018, there is a delayed response in coral health to temperature fluctuations, where the peak in coral health does not match the peak in SST. An outlier is found with the lowest coral health value of 2.55 in February 2016, which contributed to the largest decrease in coral health within the 5-year period.

4.2 Statistical Tests

According to Thompson (2019), different coral species have different optimum thermal ranges and thresholds, specific to corals found in Hong Kong reefs. The lowest temperature for a maximum thermal range is 26.4 °C, for the *Acropora samoensis* (see Figure 13). It can be assumed that coral health will decline when it is outside its optimum thermal range, possibly inducing coral bleaching. In Figures 11 and 12, the independent variable is plotted against the dependent variable as a better representation of whether there is any correlation between the two variables. Monthly SST values over 26.4 °C will be plotted. The coefficient of determination (R-squared) will be used as a statistical test as a measure of the percentage of variable variation of a linear model.

The graph in Figure 11 has an R-squared value of around 0.1133, and the graph in Figure 12 has an R-squared value of around 0.0135. The R-squared value or coefficient of determination is almost at 0% for both graphs, which indicates that the model does not explain the variability of the dependent variable around the mean. The small R-squared value on both graphs suggests that a linear regression model is not applicable for demonstrating the relationship between SST and coral health, as there is unexplainable variation possibly due to other confounding factors.

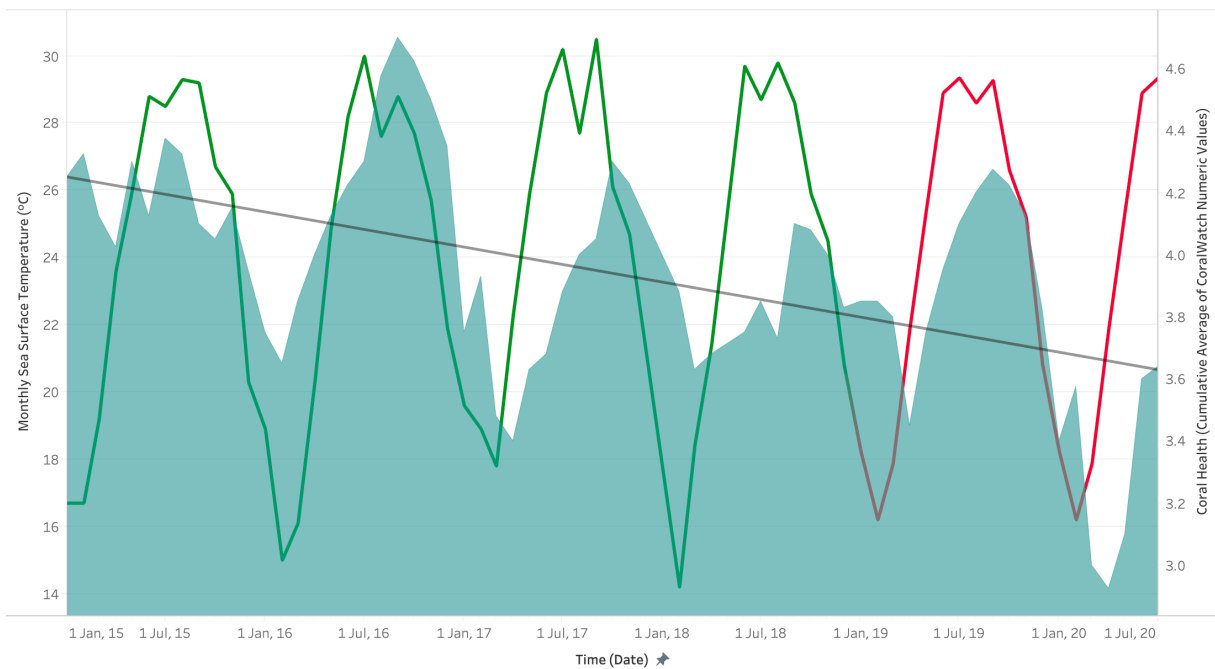


Fig. 9. Graph showing the relationship between coral health (CA) and monthly average SST ($^{\circ}\text{C}$) at Coral Beach, from 2015-2020. The green line represents monthly SST data values provided by HKO, while the red line represents extrapolated monthly SST data as no data from January 2019 to July 2020 were provided. The grey line is the trendline of the SST data set. The solid blue plot is the coral health data, which are the cumulative average of Coralwatch numeric values.

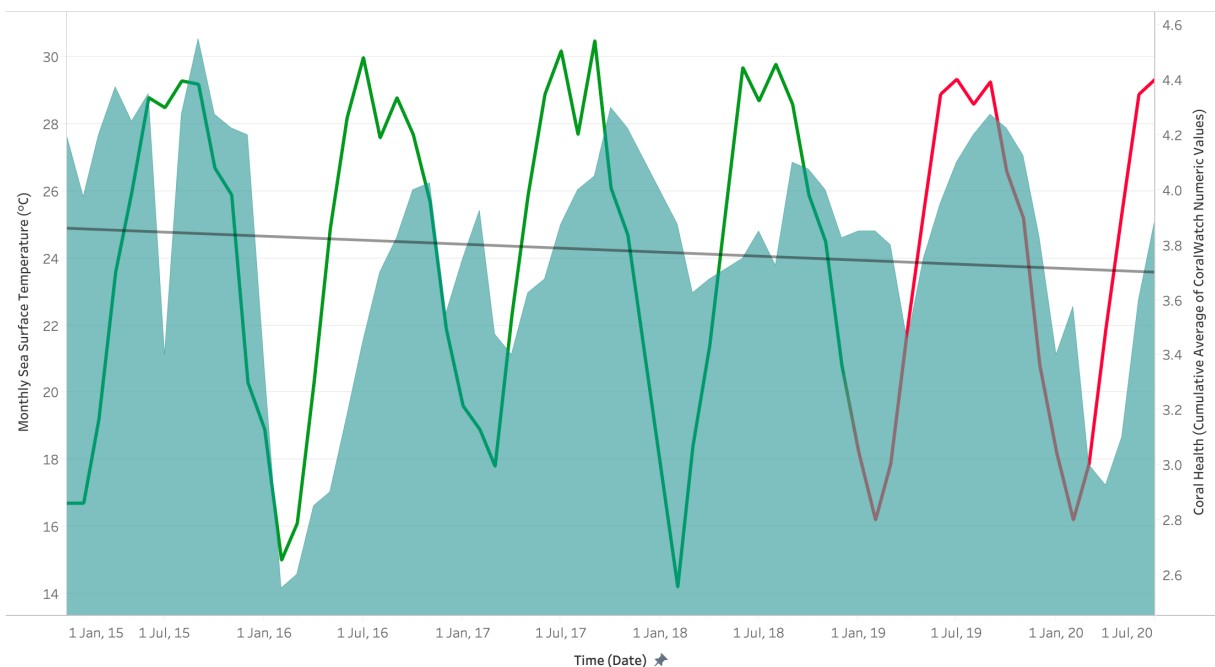


Fig. 10. Graph showing the relationship between coral health (CA) and monthly average SST ($^{\circ}\text{C}$) at the Pier Area from 2015-2020. The green line represents monthly SST data values provided by HKO, while the red line represents extrapolated monthly SST data as no data from January 2019 to July 2020 were provided. The grey line is the trendline of the SST data set. The solid blue plot is the coral health data, which are the cumulative average of Coralwatch numeric values.

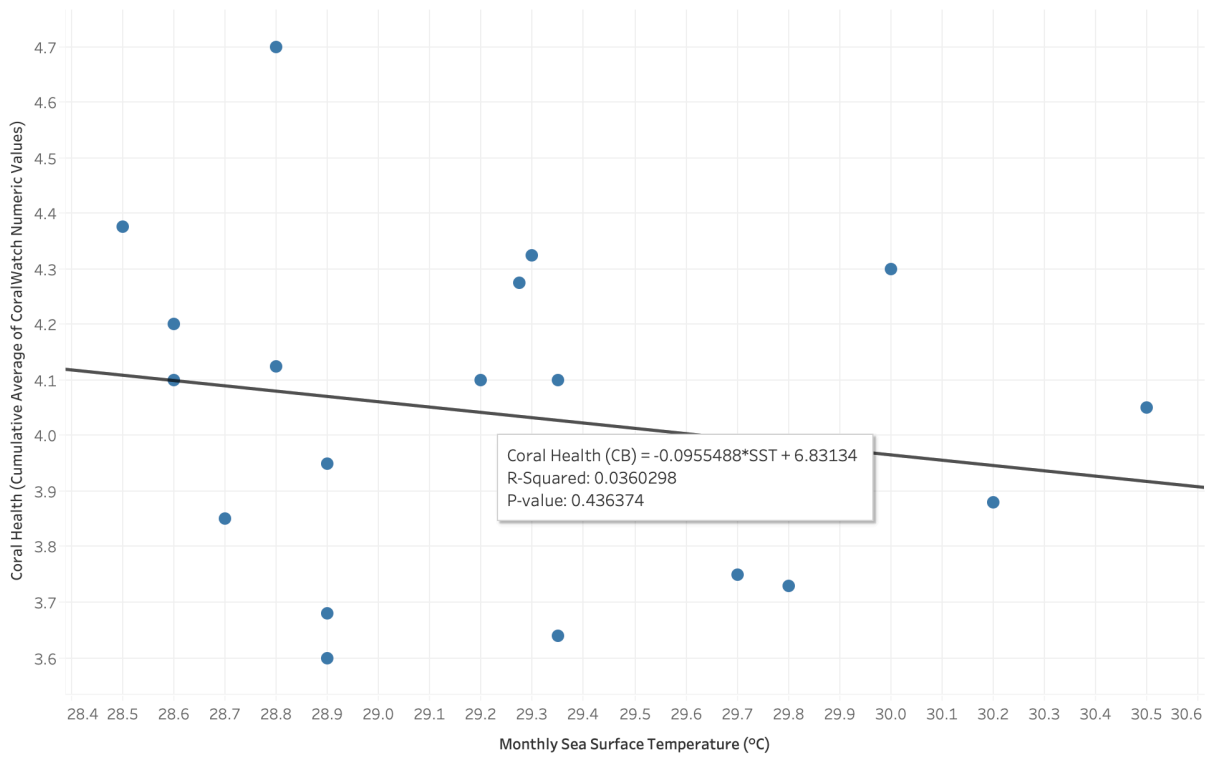


Fig. 11. Graph showing the relationship between monthly SST (above 26.4 °C) and coral health (CA) at Coral Beach.

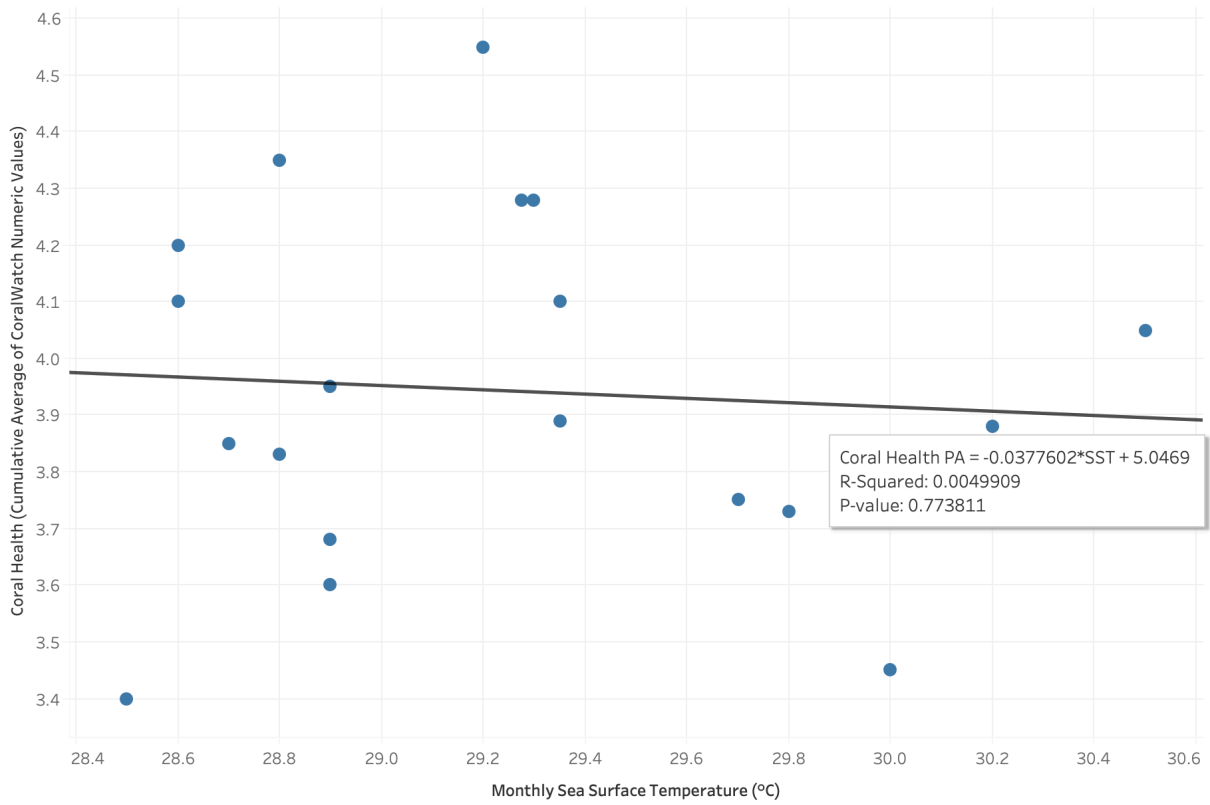


Fig. 12. Graph showing the relationship between monthly SST (above 26.4 °C) and coral health (CA) at the Pier Area.

	Optimal Thermal Range °C	Optimal Irradiance Range PAR	#Months within Optimal Conditions	#Months outside Optimal Conditions
<i>Acropora samoensis</i>	15.7 - 26.4	80 - 180	7	5
<i>Galaxea fascicularis</i>	16.9 - 27.1	80 - 175	5	7
<i>Montipora peltiformis</i>	16.1 - 27.9	90 - 240	6	6
<i>Oulastrea crispata</i>	26.5 - 32	180 - 240	3	9
<i>Porites lobata</i>	24.1 - 32	115 - 235	6	6

Fig. 13. Optimal thermal ranges for different coral species (Thompson, 2019).

5 Evaluation

5.1 Evaluation of Data

Coral health and SST fluctuate simultaneously in Figures 9 and 10. Annual decreases in coral health that happen simultaneously with the annual decreases in SST in January and February could be related to the effect of low temperatures on the metabolic rate of coral polyps while they are performing cellular respiration. Temperature has a direct effect on the functioning of enzymes involved with the tricarboxylic acid cycle (Krebs Cycle); at lower temperatures, substrates are taken up into the active sites more slowly (Thompson, 2019). Coral metabolic uptake slows down with lower temperatures, which could possibly account for lower coral health values in the winter months.

On the other hand, thermal stress can denature proteins and enzymes involved in photosynthesis, specifically, the synthesis of ATP. In photophosphorylation (Thompson, 2019), as well as enzymes involved in cellular respiration. As zooxanthellae photosynthesis has been shown to contribute up to 100% of the energy necessary for a coral's daily metabolism (Thompson, 2019), temperature stress impairing photosynthesis will hinder coral health. This may be able to explain the delayed response in the increase of coral health when temperatures increase. This is shown in Figure 9 in 2017, and in Figure 10 in 2016 to 2018 when high SST peaks at 30.0 °C and 30.5 °C respectively.

Such high temperatures could have exceeded the specific coral species thermal stress range.

The overall decreasing trend for coral health from Coral Beach may be partly explained by the abundance of *Diadema setosum* (sea urchins) at Coral Beach. Lam *et al.*, (2007) reported that an outbreak of *D. setosum* caused severe bioerosion, weakened the CaCO₃ skeleton of corals making them easy to topple and vulnerable to diseases (A. L. Chan *et al.*, 2005), in a large number of coral colonies in Coral Beach in 2006-2007. Coral colonies at Moon Island, also in Hoi Ha Wan have been completely bio-eroded (with a coral cover of only 2%, as of June, 2020), according to Paul Hodgson from Oceanway Corporation Limited, based in Hoi Ha Wan.

Despite the low R-squared value, it is worth noting there is an obvious gradient on the linear regression line for Figure 11 (-0.103) which suggests a weak inverse relationship as coral health decreases while SST increases above 26.4 °C in Coral Beach.

5.2 Evaluation of procedure

Usually, when studying the effects of sea temperature changes on coral health, researchers conduct experiments in laboratory settings, as demonstrated by Thompson (2019), Guan *et al.* (2015), Jokiel (2004). This allows for the isolation of the independent variable, unlike *in situ* studies, which had uncontrollable confounding variables. How-

ever, monitoring real-life situations is a realistic and essential measure.

Middlebrook *et al.* (2008), reported that bleaching can occur within the span of 5 days (if waters are heated to 34 °C). Therefore, collecting SST data once per month is not representative of any sudden anomalies or fluctuations that could occur within a day, and the use of a monthly average would also smooth out extreme and drastic temperature variations occurring within days, or hours.

Thompson (2019) has demonstrated that different species of corals (most commonly found in Hong Kong) have different optimum thermal ranges, the amount of time spent outside their optimum ranges affects coral health, and tolerate different levels of exposure to heat stress (Thompson, 2019). When conducting the line transect, the type of coral (branching, boulder, plate or soft) was recorded as an observation, but was not used as part of the data analysis. The transect lines are not fixed, so each dive will sample different proportions of different coral species, which have varying levels of tolerance to thermal stress, thus affecting the overall coral reef health (CA value).

The use of a line transect survey ensures that no same coral is sampled more than once. However, even with the use of an exact GPS coordinate, the location where the line transect is placed varies for each dive. Therefore, different sections of the reef are sampled for each dive, which may include varying degrees of coral health and coral species proportions.

6 Conclusion

This exploration aimed to investigate how SST changes above coral tolerance ranges would affect coral health in Hoi Ha Wan, from 2015-2020. The study is limited by the lack of detailed information on short-term SST fluctuations, and the inaccuracy in the coral sampling method for monitoring coral health. Whilst this study did not confirm that temperature rises induce coral bleaching events, it illustrated that coral health fluctuates with annual SST variations. The linear regression line in Figure 11 substantiates the hypothesis (H_1) by show-

ing a weak inverse relationship where coral health declines as monthly SST increases above the optimum thermal limit.

While coral bleaching is caused by multiple factors, the relationship between temperature and coral health has been established, coral bleaching is triggered when there are 1-2 °C above-average SST anomalies (Marshall *et al.*, 2006). SST data collected once per month cannot reflect the SST fluctuations that could occur in the span of a few hours or days. The precise mechanism of changes in short-term SST on coral health, specifically for corals in Hoi Ha Wan remains to be elucidated. More frequent data collection on SST fluctuations would establish a greater degree of accuracy. Collection of hourly SST data, by *in situ* data loggers, can be used to identify and establish the causes of coral bleaching incidents, or even identify coral acclimatization and adaptations to withstand temperature changes outside their temperature tolerance range (David *et al.*, 2013).

Increased participation from recreational divers and citizen scientists could be encouraged for greater amounts of coral health data. Permanent transects could be used as a data collection method to include sufficient amounts of different coral species to study their different responses to thermal stress and would allow for the monitoring of the same corals over time (Salmond *et al.*, 2017).

If global and local threats to corals are left unchecked, it could result in mass coral mortality in the future with huge impact on biodiverse ecosystems affecting enormous ocean life and the human race. As reefs in Hong Kong survive among the highest SST range in the world and are more stress-resistant than corals in tropical climates (Goodkin *et al.*, 2011), corals from Hong Kong are an excellent model for studying how corals cope with the changes in climate.

References

- Bradbeer, R., Hodgson, P., Ku, K. K., Lam, K., Yeung, L. F., & Zhan, C. J. (2008). Towards a 3-dimensional model of Hoi Ha Wan Marine Park, Hong Kong. *OCEANS 2008*. doi: 10.1109/oceans.2008.5152055
- Chan, A., & Hodgson, P. (2018). Improving Stakeholder and Authority Cooperation Among Coastal Fishing Communities Using Passive Blast Monitoring Data. *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*. doi: 10.1109/oceanskobe.2018.8559100
- Chan, A. L., Choi, C. L., McCorry, D., Chan, K. K., Lee, M., & Put, A. (2005). *Field Guide to Hard Corals of Hong Kong*. Cosmos Books and Friends of the Country Parks.
- Cope, M. (1986). Seasonal, diel and tidal hydrographic patterns, with particular reference to dissolved oxygen, above a coral community at Hoi Ha Wan, Hong Kong. *Asian Marine Biology*, 3, 59-74.
- Cope, M., & Morton, B. (1988). The Scleractinian Coral Community At Hoi Ha Wan, Hong Kong. *Asian Marine Biology*, 5, 41-52.
- Cornish, A. (2003). *Diadema Sea Urchins and the Black-Spot Tuskfish*. Department of Ecology and Biodiversity, the University of Hong Kong. Retrieved 3 September, 2020, from <https://www.biosch.hku.hk/ecology/porcupine/por28/28-vert-urchin.htm>
- David, A., Atkinson, A., Feeley, M., Miller, J., Patterson, J., Richter, L., ... Witcher, B. (2013). *Coral Reef Ecosystem Water Temperature Monitoring Protocol*. National Park Service.
- Environmental Protection Department. (2020). *Marine Water Quality Monitoring Data*. Environmental Protection Department. Retrieved 3 September, 2020, from <https://cd.epic.epd.gov.hk/EPICRIVER/marine/>
- Goodkin, N. F., Switzer, A. D., McCorry, D., Devantier, L., True, J. D., Hughen, K. A., ... Yang, T. T. (2011). Coral communities of Hong Kong: long-lived corals in a marginal reef environment. *Marine Ecology Progress Series*, 426, 185-196.
- Guan, Y., Hohn, S., & Merico, A. (2015). Suitable Environmental Ranges for Potential Coral Reef Habitats in the Tropical Ocean. *PLOS ONE*, 10, e0128831. doi: 10.1371/journal.pone.0128831
- Hoegh-Guldberg, O., & Lough, J. (2015). *Climate change and coral bleaching*. Australian Academy of Science. Retrieved 3 September, 2020, from <https://www.science.org.au/curious/coral-bleaching>
- Hoegh-Guldberg, O., Pendleton, L., & Kaup, A. (2019). People and the changing nature of coral reefs. *Regional Studies in Marine Science*, 30, 100699. doi: 10.1016/j.rsma.2019.100699
- Holland, G. J. (2009). Predicting El Niño's Impacts. *Science*, 325, 47-47. <https://science.sciencemag.org/content/325/5936/47.full?rss=1> doi: 10.1126/science.1176515
- Hong Kong Observatory. (2019). *Monthly Mean Sea Temperature (°C) (a.m.) at North Point*. Hong Kong Observatory. Retrieved 3 September, 2020, from http://www.hko.gov.hk/en/cis/monthlyElement.htm?ele=SEATEMP_NP_AM
- Jokiel, P. L. (2004). Temperature Stress and Coral Bleaching. *Coral Health and Disease*, 401-425. doi: 10.1007/978-3-662-06414-6_23
- Lam, K., Shin, P. K. S., & Hodgson, P. (2007). Severe bioerosion caused by an outbreak of corallivorous *Drupella* and *Diadema* at Hoi Ha Wan Marine Park, Hong Kong. *Coral Reefs*, 26, 893-893. doi: 10.1007/s00338-007-0288-9
- Marshall, P., Schuttenberg, H., West, J., & Great Barrier Reef Marine Park Authority. (2006). *A reef manager's guide to coral bleaching*. Great Barrier Reef Marine Park Authority.
- Middlebrook, R., Hoegh-Guldberg, O., & Leggat, W. (2008). The effect of thermal history on the susceptibility of reef-building corals to thermal stress. *Journal of Experimental Biology*, 211, 1050-1056. doi: 10.1242/jeb.013284
- Orlowski, J., & Rhodes, L. (2017). *Chasing Coral*. Retrieved 2020-09-03, from <https://www.youtube.com/watch?v=aGGBGcjdjXA>
- Salmond, J., Loder, J., Roelfsema, C., & Passenger, J. (2017). *Reef Check Australia 2017 Heron Reef Health Report*. Reef Check Australia.
- Schmidt, C. W. (2008). In Hot Water: Global Warming Takes a Toll on Coral Reefs. *Environmental Health Perspectives*, 116. doi: 10.1289/ehp.116-a292
- Tameland, J. (2016). *Coral Reefs - Valuable but Vulnerable*. United Nations Environment Programme. Retrieved 3 September, 2020, from http://coral.unep.ch/Coral_Reefs.html
- Thompson, P. D. (2019). *Physiological responses to seasonality in five species of subtropical coral* (Doctoral dissertation). Retrieved 2020-09-03, from <https://hub.hku.hk/handle/10722/281274>
- Thornhill, D. (2012). *Ecological Impacts and Practices of the Coral Reef Wildlife Trade*. Defenders of Wildlife.
- United Nations et al. (2012). *Geo 5: Global Envi-*

ronment Outlook: Environment for the future we want. United Nations Environment Program.

Veron, J., Stafford-Smith, M., Turak, E., & DeVantier, L. (2016). *Coral structure and growth.* Retrieved 3 September, 2020, from <http://www.coralsoftheworld.org/page/structure-and-growth/>

Wong, A. (2006). *20 years of Marine Water Quality Monitoring in Hong Kong.* Environmental Protection Department. Retrieved 3 September, 2020, from https://www.epd.gov.hk/epd/misc/marine_quality/1986-2005/eng/03_development_menu.htm

A study on the relationship between the colour and mass of selected main sequence stars

Candace Yan Yue Chung 鍾欣瑜

1 Research Question

How is the mass and colour of selected main sequence stars related?

2 Introduction

The mass of a star determines the star's temperature, luminosity, stability, and gravity, all of which are important indicators of the evolution of a star (Torres *et al.*, 2009). As it is impossible to directly measure the mass of the star, scientists must innovate different methods to indirectly measure stellar mass. Even though numerous methods used to measure stellar mass, such as the velocity method (Kepler, 1968), involve the observation in the change of the star's colours, there is no equation in the textbook that directly relates the wavelength of the star's colour to the stellar mass. Therefore, I decided to derive the colour-mass relationship for stars and verify its accuracy.

3 Background

3.1 Main Sequence Stars

Main-sequence stars are stars that are powered by hydrogen fusion (Unsöld & Baschek, 1983). On the H-R diagram, which is a diagram where the wavelength of light emitted by stars are plotted against their luminosity, the main sequence stars appear on the diagram as a distinct band (Russel, 1914).

3.2 Important formulae and relationships

3.2.1 Colour and temperature

The relationship between the colour and temperature of a star can be described using Wien's displacement law (1), where λ_{max} is the wavelength

of peak radiation emission of the star in metres, b is the Wien constant 2.9×10^{-3} mK, and T is the temperature in Kelvin. This formula shows that the wavelength of the colour of a star is inversely proportional to its temperature.

$$\lambda_{max} = \frac{b}{T} \quad (1)$$

3.2.2 Luminosity and temperature:

The Stefan-Boltzmann law (2) describes the relationship between the luminosity and temperature of a star, where L is the luminosity of a body, e is the emissivity of the body which equals 1 for black bodies and stars, σ is the Stefan-Boltzmann constant ($5.76 \times 10^{-8} \text{Wm}^{-2}\text{K}^{-4}$), and A is the surface area of the body. This law shows that luminosity of the star is related to temperature raised to the fourth power.

$$L = e\sigma AT^4 \quad (2)$$

3.2.3 Mass and luminosity:

The more massive a star is, the greater its gravity. As a result, stars would need to produce greater radiation pressure to support the star and prevent it from collapsing into itself (Salaris & Cassisi, 2008). Even though this source is published in 2008, this is one of the most recent papers about the mass-luminosity relationships, which means that the information this source provides should still be accurate. Although in general the mass of a star is related to its luminosity, the specific mass-luminosity relationship differs according to different mass ranges (Duric, 2004). In the IB physics textbooks, the mass-luminosity relationship is stated as follows (Tsokos, 2014):

$$\frac{L_*}{L_{\odot}} = \left(\frac{M_*}{M_{\odot}}\right)^{3.5} \quad (3)$$

However, (3) only applies to main-sequence stars 2 to 55 times the mass of the sun (Lecchini, 2007). For main-sequence stars 0.43 to 2 times the mass of the sun, the mass-luminosity relationship is (4), where L_* is the luminosity of a star, M_* is the mass of a star, L_\odot is the luminosity of the sun, and M_\odot is the mass of the sun (Salaris & Cassisi, 2008).

$$\frac{L_*}{L_\odot} = \left(\frac{M_*}{M_\odot}\right)^4 \quad (4)$$

3.3 Common methods used to indirectly measure stellar mass

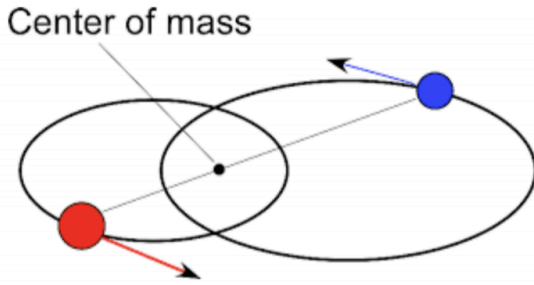


Fig. 1. Diagram of two stars orbiting each other in a binary star system (Gowdy, 2009)

Binary star systems are systems where two stars orbit each other around a centre of mass (Fig. 1)). If the star exists in a binary/multi-star system, the star's velocity can be used to calculate their stellar mass (Kepler, 1968). As stars orbit each other, the wavelengths of the radiation emitted by the stars will change from being separate wavelengths to being a single superimposed wavelength, depending on their positions. From these changes in wavelengths, the Doppler shift in the stars' velocities can be observed, and the stellar masses can be calculated using Kepler's third law (5), where t is the orbital period of stars, a is the semi-major axis of the star's orbit, G is the gravitational constant $6.67 \cdot 10^{-11}$ N, M_1 is the mass of the more massive object, and M_2 is the mass of the less massive object (Kepler, 1968).

$$t^2 = \frac{4\pi^2 a^3}{G(M_1 + M_2)} \quad (5)$$

4 Colour-mass relationship derivation

4.1 Colour-mass relationship derivation assumptions

1. Stars are in main-sequence
2. Stars are spherical masses to allow for easier approximation of the star's surface area
3. Stars are emitting light at the wavelength of peak radiation emission λ_{max}
4. Radiant energy is absorbed and released at the same rate
5. Stars radiate energy characteristic of the temperature
6. Stars behave like perfect black bodies (emissivity equals 1)

As the mass-luminosity relationship is different for stars between 2 to 55 times the mass of the sun and for stars between 0.43 to 2 times the mass of the sun, the colour-mass relationship derivations should be separated into these two categories.

4.2 For stars with masses between 2 to 55 times the mass of the sun

Rearrange the mass-luminosity formula (3) so L_* is the subject

$$L_* = L_\odot \left(\frac{M_*}{M_\odot}\right)^{3.5} \quad (6)$$

Substitute the above formula into the Stefan-Boltzmann law (2)

$$L_\odot \left(\frac{M_*}{M_\odot}\right)^{3.5} = e\sigma AT^4 \quad (7)$$

Rearrange Wien's displacement law (1) so T is the subject

$$T = \frac{b}{\lambda_{max}} \quad (8)$$

Substitute (8) back into (7) and replace A in (7) with surface area of a sphere ($A = 4\pi r^2$)

$$L_\odot \left(\frac{M_*}{M_\odot}\right)^{3.5} = 4\pi r^2 e\sigma \left(\frac{b}{\lambda_{max}}\right)^4 \quad (9)$$

Isolate M_* and take into account $e = 1$ for stars

$$M_* = \sqrt[3.5]{4\pi r^2 \sigma \left(\frac{M_\odot}{L_\odot}\right)^{3.5} \left(\frac{b}{\lambda_{max}}\right)^4} \quad (10)$$

From this derivation, it can be concluded that for stars with a mass between 2 to 55 times that of the sun, $M_* \propto \lambda_{max}^{-\frac{4}{3.5}}$.

4.3 For stars with masses between 0.43 to 2 times the mass of the sun

Repeat the derivation for stars with masses between 0.43 to 2 times the mass of the sun but instead, start with the mass-luminosity formula (4). Doing so produces the following expression:

$$M_* = \sqrt[4]{4\pi r^2 \sigma \left(\frac{M_\odot}{L_\odot}\right)^4 \left(\frac{b}{\lambda_{max}}\right)^4} \quad (11)$$

From this derivation, it can be concluded that for stars with a mass between 0.43 to 2 times the mass of the sun, stellar mass should also be inversely proportional to its wavelength of colour emitted.

5 Variables for analysis

As the derived formulas involve three variables (stellar mass, stellar radius, and wavelength of colour), the radius and wavelength will be combined as one variable to simplify analysis.

5.1 For stars with masses between 2 to 55 times the mass of the sun

The colour-mass relationship for stellar masses between 2-55 times the mass of the sun is (10):

$$M_* = \sqrt[3.5]{4\pi r^2 \sigma \left(\frac{M_\odot}{L_\odot}\right)^{3.5} \left(\frac{b}{\lambda_{max}}\right)^4} \quad (12)$$

The colour-mass relationship can then be rearranged as so:

$$M_* = \sqrt[3.5]{4\pi \sigma \frac{M_\odot^{3.5}}{L_\odot}} b^4 \times \sqrt[3.5]{\frac{r^2}{\lambda_{max}^4}} \quad (13)$$

The combined variable will then be this:

$$\sqrt[3.5]{\frac{r^2}{\lambda_{max}^4}} \quad (14)$$

Therefore, M_* should be directly proportional to $\sqrt[3.5]{\frac{r^2}{\lambda_{max}^4}}$. As $\sqrt[3.5]{4\pi \sigma \frac{M_\odot^{3.5}}{L_\odot}} b^4$ is a constant, the mathematical structure of (13) implies that $\sqrt[3.5]{4\pi \sigma \frac{M_\odot^{3.5}}{L_\odot}} b^4$ acts as the slope of (13).

5.2 For stars with masses between 0.43 to 2 times the mass of the sun

The colour-mass relationship for stellar masses between 0.43-2 times the mass of the sun is (11):

$$M_* = \sqrt[4]{4\pi r^2 \sigma \left(\frac{M_\odot}{L_\odot}\right)^4 \left(\frac{b}{\lambda_{max}}\right)^4} \quad (15)$$

The colour-mass relationship can then be rearranged as so:

$$M_* = \sqrt[4]{4\pi \sigma \frac{M_\odot^4}{L_\odot}} b^4 \times \frac{\sqrt{r}}{\lambda_{max}} \quad (16)$$

The combined variable will then be this:

$$\frac{\sqrt{r}}{\lambda_{max}} \quad (17)$$

Therefore, M_* should be directly proportional to $\frac{\sqrt{r}}{\lambda_{max}}$. As $\sqrt[4]{4\pi \sigma \frac{M_\odot^4}{L_\odot}} b^4$ is a constant, the mathematical structure of (16) implies that $\sqrt[4]{4\pi \sigma \frac{M_\odot^4}{L_\odot}} b^4$ acts as the slope of (16).

6 Methodology

6.1 Database selection

6.1.1 Mass and radii values

As stellar mass is indirectly measured, many stellar databases do not include the mass and radii values of stars. Torres' database was selected not only

because of the meticulous verification of the accuracy of mass and radii values through the systematic recomputation of these values but also the inclusion of absolute uncertainty values for mass, which is not included in many other databases.

6.1.2 Miscellaneous stellar data

The HYG Database was selected for the source of other data values because it has a large collection of stars (119615 stars in total) and a wide variety of detailed information such as the B-V colour indices (Nash, 2006) of stars, which is information that is required to substitute into the colour-mass model.

6.2 Star selection

6.2.1 Star selection criteria

1. Stars have to appear in both datasets (so no star has any missing data and to allow for easier comparison with the database mass values for the evaluation of the methods)
2. Stars have to be main-sequence stars (so the formulas used are applicable to them)

As there were more stars in the HYG database, the stars present in Torres' database were searched in the HYG database to see which stars are present in both databases. Afterwards, the stars had to be confirmed to be main-sequence stars. This is important as all the existing formulas only apply to stars in the main-sequence.

6.2.2 Main-sequence star verification

Conversion from the colour index (B-V index) to wavelength

The B-V colour index is a number that describes the colour of an object (Dintsis, Artemi, & Polatoglou, 2018). The more negative the B-V colour index value is, the shorter the wavelength of light the star emits (Dintsis et al., 2018). For example, Star HD 86118 has a B-V index of -0.173, indicating that it is emitting light with a wavelength of approximately 224 nanometers (Nash, 2006). As the information regarding the wavelength of light emitted by each star is in the form of B-V colour indices, the B-V colour indices need to be converted

to wavelength for ease of graphing the HR diagram and performing other calculations. The following formula converts the B-V colour index to temperature (K), where I_{B-V} represents the index value (Dintsis et al., 2018):

$$T = 4600 \left(\frac{1}{0.92(I_{B-V}) + 1.7} + \frac{1}{0.92(I_{B-V}) + 0.62} \right) \quad (18)$$

Substitute the rearranged Wien's displacement law with T as the subject (8) into (18) and rearrange so λ_{max} is the subject

$$\lambda_{max} = \frac{2.9 \times 10^{-3}}{4600 \left(\frac{1}{0.92(I_{B-V}) + 1.7} + \frac{1}{0.92(I_{B-V}) + 0.62} \right)} \quad (19)$$

After converting the B-V index to wavelength, the wavelength of light emitted by each star was graphed against the luminosity values of the stars, in which the scale of the luminosities is on a logarithmic scale. As seen from Fig. 2, mass values of all the stars approximately lie in a straight, distinct band, indicating that they can be assumed to be main-sequence stars. Of the 119710 stars available in both databases, 31 stars fulfill both criteria.

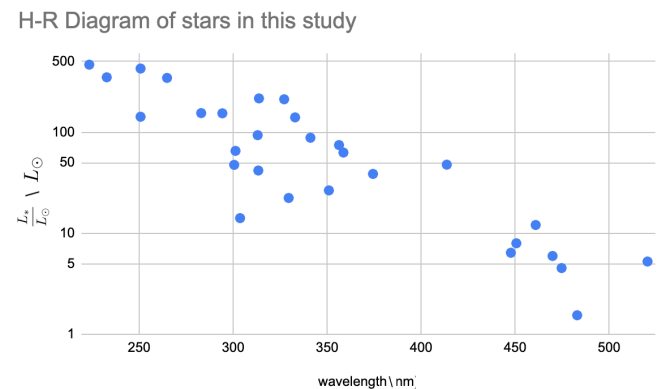


Fig. 2. H-R diagram of the 31 stars included in the study

7 Data Processing

7.1 Calculating mass using colour-mass models

In order to graph the colour-mass relationship against the database mass values, the stellar mass needs to be calculated using the colour-mass models. Since the mass-luminosity relationship varies between different mass ranges, it is worthwhile to

compute the mass values of all the stars using both variations of the colour-mass formula.

Star HD165921 sample calculation (using the formula for masses between 0.43 to 2 times the mass of the sun)

$$M_* = \sqrt[4]{4\pi(4.265 \times 10^9)^2\sigma\left(\frac{M_\odot}{L_\odot}\right)}\left(\frac{b}{3.138 \times 10^{-7}}\right) \quad (20)$$

$$M_* = 1.294 \times 10^{27} \text{ kg} \quad (21)$$

Data

Henry Draper ID	Wavelength (10 ⁷ m)	Radius (10 ⁹ m)	Calculated Mass (10 ²⁷ kg)
165921	3.138	4.265	1.294
86118	2.235	2.822	1.477
42401	2.652	2.618	1.199

Table 1. Calculated using the formula for masses between 0.43 to 2 times the mass of the sun, first 3 values displayed as examples.

Star HD165921 sample calculation (using the formula for masses between 2 to 55 times the mass of the sun)

$$M_* = \sqrt[3.5]{4\pi(4.265 \times 10^9)^2\sigma\left(\frac{M_\odot}{L_\odot}\right)}\left(\frac{b}{3.138 \times 10^{-7}}\right)^4 \quad (22)$$

$$M_* = 9.658 \times 10^{30} \text{ kg} \quad (23)$$

Data

Henry Draper ID	Wavelength (10 ⁷ m)	Radius (10 ⁹ m)	Calculated Mass (10 ³⁰ kg)
165921	3.138	4.265	9.658
86118	2.235	2.822	11.24
42401	2.652	2.618	8.856

Table 2. Calculated using the formula for masses between 2 to 55 times the mass of the sun, first 3 values displayed as examples.

7.2 Measuring uncertainty

7.2.1 Wavelength relative uncertainty

As the HYG database doesn't provide wavelength values but rather uses the B-V colour index, the measurement uncertainty values for wavelength will be derived from the wavelength to B-V colour

index equation (19). As (19) shows the B-V colour index in a fraction, the measurement uncertainty for the wavelength of colour could be approximated as following:

$$\frac{\Delta\lambda_{max}}{\lambda_{max}} = \frac{\Delta I_{B-V}}{I_{B-V}} + \frac{\Delta I_{B-V}}{I_{B-V}} \quad (24)$$

$$\frac{\Delta\lambda_{max}}{\lambda_{max}} = 2\frac{\Delta I_{B-V}}{I_{B-V}} \quad (25)$$

Since the B-V colour index listed in the HYG database does not include measurement uncertainty, the measurement uncertainty in the B-V colour index will be assumed to be half of the last significant figure. As the B-V colour index is provided up to 3 decimal places in the database, the absolute uncertainty is ± 0.0005 . Additionally, as the B-V colour index allows for negative values, the colour index values used in uncertainty calculations will be its absolute value.

Star HD165921 sample calculation

$$\frac{\Delta I_{B-V}}{I_{B-V}} = \frac{0.0005}{0.0790} \quad (26)$$

$$\frac{\Delta I_{B-V}}{I_{B-V}} = 6.329 \times 10^{-3} \quad (27)$$

$$\frac{\Delta\lambda_{max}}{\lambda_{max}} = 2 \times 6.329 \times 10^{-3} \quad (28)$$

$$\frac{\Delta\lambda_{max}}{\lambda_{max}} = 1.266 \times 10^{-2} \quad (29)$$

Data

Henry Draper ID	B-V Relative Uncertainty (10 ⁻³)	Wavelength Relative Uncertainty (10 ⁻³ m)
165921	6.329	12.66
86118	2.890	5.780
42401	8.475	16.95

Table 3. First 3 values displayed as examples.

7.2.2 Radius relative uncertainty

The absolute uncertainty values for stellar radius are listed up to 3 decimal places in Torres' database. However, the radii were given in solar radii (R_{solar}). Therefore, the units need to be con-

verted from solar radius (R_{solar}) to metres. The unit conversion equation is as follows:

$$r = R_{solar} \times 696340 \times 10^3 \quad (30)$$

After the conversion of units, the radius relative uncertainty is simply $\frac{\Delta r}{r}$.

Star HD165921 sample calculation

Absolute uncertainty unit conversion:

$$r = 0.060 \times 696340 \times 10^3 \quad (31)$$

$$r = 4.178 \times 10^7 \quad (32)$$

Radius unit conversion:

$$r = 6.13 \times 696340 \times 10^3 \quad (33)$$

$$r = 4.265 \times 10^9 \quad (34)$$

Radius relative uncertainty calculation:

$$\frac{\Delta r}{r} = \frac{4.178 \times 10^7}{4.265 \times 10^9} \quad (35)$$

$$\frac{\Delta r}{r} = 9.796 \times 10^{-3} \quad (36)$$

Data

Henry Draper ID	Radius (10^9 m)	Absolute Uncertainty (10^7 m)	Relative Uncertainty (10^{-3} m)
165921	4.265	4.178	9.769
86118	2.822	6.337	22.45
42401	2.618	4.178	15.96

Table 4. First 3 values displayed as examples.

7.2.3 Calculation of calculated mass uncertainty from colour-mass relationship (for the colour-mass relationship describing stars 2-55 times the mass of the sun)

The colour-mass relationship for stellar masses (M_*) between 2-55 times the mass of the sun is (10).

$$M_* = \sqrt[3.5]{4\pi r^2 \sigma \left(\frac{M_\odot}{L_\odot}\right)^{3.5} \left(\frac{b}{\lambda_{max}}\right)^4} \quad (37)$$

Assuming the mass and luminosity of the sun

Data

Henry Draper ID	Radius (Relative) (10^{-3} m)	Wavelength (Relative) (10^{-3} m)	Mass (Relative) (10^{-2} m)
165921	6.329	12.66	2.006
86118	2.890	5.780	1.944
42401	8.475	1.695	2.894

Table 5. First 3 values displayed as examples.

As seen from these calculations, the uncertainty of the calculated mass values for the colour-mass relationship for stars 2-55 times the mass of the sun is very small, indicating that this colour-mass relationship produces very precise calculated mass values.

7.2.4 Calculation of calculated mass uncertainty from colour-mass relationship (for the colour-mass relationship describing stars 0.43-2 times the mass of the sun)

The colour-mass relationship for stellar masses between 0.43-2 times the mass of the sun is (11).

$$M_* = \sqrt[4]{4\pi\sigma \frac{M_\odot^4}{L_\odot} b^4} \times \frac{\sqrt{r}}{\lambda_{max}} \quad (41)$$

Assuming the mass and luminosity of the sun are constants, the calculated mass uncertainty derived from (11) is as follows:

$$\frac{\Delta M_*}{M_*} = \frac{1}{2} \frac{\Delta r}{r} + \frac{\Delta \lambda_{max}}{\lambda_{max}} \quad (42)$$

Star HD165921 sample calculation

$$\frac{\Delta M_*}{M_*} = \frac{1}{2} \times 9.796 \times 10^{-3} + 1.266 \times 10^{-2} \quad (43)$$

$$\frac{\Delta M_*}{M_*} = 1.756 \times 10^{-2} \quad (44)$$

Data

As seen from these calculations, it is clear that the colour-mass relationship describing stars 0.43-2 times the mass of the sun is also very small, indicating that this colour-mass relationship produces very precise calculated mass values.

Henry Draper ID	Radius (Relative) (10^3 m)	Wavelength (Relative) (10^3 m)	Mass (Relative) (10^3 m)
165921	6.329	12.66	1.756
86118	2.890	5.780	1.701
42401	8.475	1.695	2.493

Table 6. First 3 values displayed as examples.

7.2.5 Database stellar mass values uncertainty:

Torres' database stellar mass values (m_*) also includes absolute uncertainty in solar masses up to 3 decimal places (M_{solar}). Therefore, the units need to be converted from solar masses (M_{solar}) to kg. The unit conversion equation is as follows:

$$m_* = M_{solar} \times 1.989 \times 10^{30} \quad (45)$$

Star HD165921 sample calculation

Absolute uncertainty unit conversion:

$$m_* = 0.440 \times 1.989 \times 10^{30} \quad (46)$$

$$m_* = 8.752 \times 10^{29} \quad (47)$$

Mass unit conversion:

$$r = 19.01 \times 1.989 \times 10^{30} \quad (48)$$

$$r = 3.781 \times 10^{31} \quad (49)$$

Mass relative uncertainty calculation:

$$\frac{\Delta m_*}{m_*} = \frac{8.752 \times 10^{29}}{3.781 \times 10^{31}} \quad (50)$$

$$\frac{\Delta m_*}{m_*} = 2.315 \times 10^{-2} \quad (51)$$

Data

Henry Draper ID	Mass (10^{31} kg)	Absolute Uncertainty (10^{29} m)	Relative Uncertainty (10^2 m)
165921	3.781	8.752	2.315
86118	1.683	2.387	1.418
42401	1.026	1.193	1.163

Table 7. First 3 values displayed as examples.

8 Analysis

8.1 Computation of combined variables

The combined variables refer to (14) and (17). These variables will be graphed on the x-axis of the graphs.

8.1.1 For the colour-mass relationship describing stars 2-55 times the mass of the sun

Star HD165921 sample calculation

$$\sqrt[3.5]{\frac{r^2}{\lambda_{max}^4}} = \sqrt[3.5]{\frac{(4.265 \times 10^9)^2}{(3.138 \times 10^{-7})^4}} \quad (52)$$

$$\sqrt[3.5]{\frac{r^2}{\lambda_{max}^4}} = 8.616 \times 10^{12} \quad (53)$$

Data

Henry Draper ID	Radius (10^9 m)	Wavelength (10^{-7} m)	Calculated Variable (10^{12} m)
165921	4.265	3.138	8.616
86118	2.822	2.235	10.03
42401	2.618	2.652	7.900

Table 8. First 3 values displayed as examples.

8.1.2 For the colour-mass relationship describing stars 0.43-2 times the mass of the sun

Star HD165921 sample calculation

$$\frac{\sqrt{r}}{\lambda_{max}} = \frac{\sqrt{4.265 \times 10^9}}{3.1380 \times 10^{-7}} \quad (54)$$

$$\frac{\sqrt{r}}{\lambda_{max}} = 2.081 \times 10^{11} \quad (55)$$

Data

Henry Draper ID	Radius (10^9 m)	Wavelength (10^{-7} m)	Calculated Variable (10^{11} m)
165921	4.265	3.138	2.081
86118	2.822	2.235	2.377
42401	2.618	2.652	1.929

Table 9. First 3 values displayed as examples.

8.2 Graphs

8.2.1 For the colour-mass relationship describing stars 2-55 times the mass of the sun

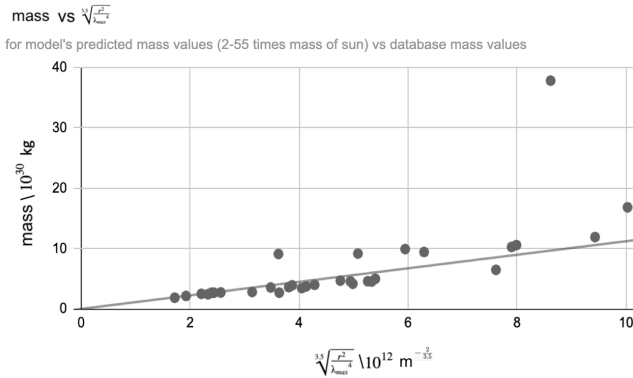


Fig. 3. Graph comparing database mass values with the model for stellar masses 2-55 times the mass of the sun (including outlier)

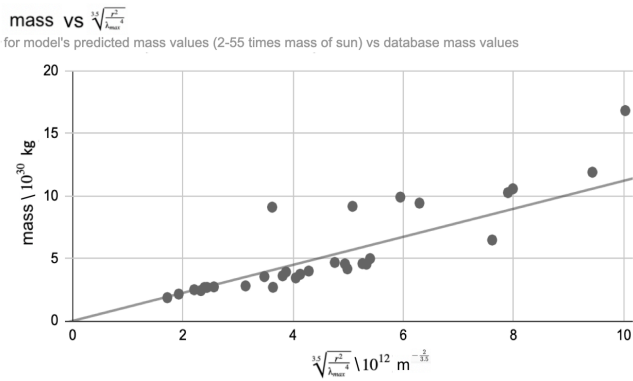


Fig. 4. Graph comparing database mass values with the model for stellar masses 2-55 times the mass of the sun (without outlier)

The grey points represent database mass values (Fig. 3 and 4). The grey trendline is the model function (Fig. 3 and 4). The error bars in both graphs are so small they are covered by the points themselves (Fig. 3 and 4). In general, the points display a positive linear trend. Stars with masses between 1.856×10^{30} to 4.981×10^{30} kg generally do not deviate from the function greatly, with stars between 1.856×10^{30} to 2.721×10^{30} kg almost completely aligning with the function (Fig. 3 and 4). The star with the mass 3.781×10^{31} kg is an obvious outlier (Fig. 3). Despite the fact that the colour-mass function is supposed to describe stars with higher masses (2-55 times the mass of the sun), the data points of stars with greater masses actually deviate from the function more.

8.2.2 For the colour-mass relationship describing stars 0.43-2 times the mass of the sun

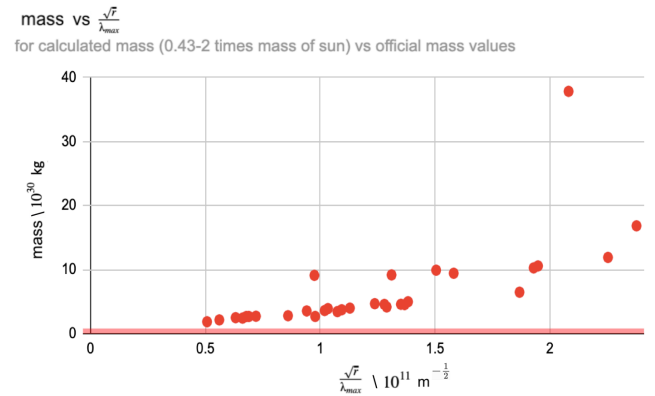


Fig. 5. Graph comparing database mass values with the model for stellar masses 0.43-2 times the mass of the sun (with outlier)

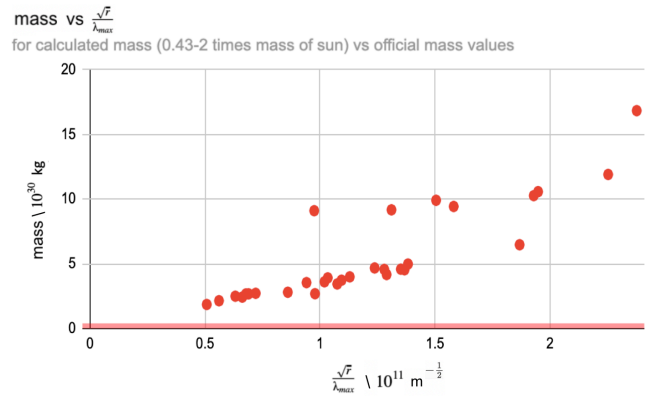


Fig. 6. Graph comparing database mass values with the model for stellar masses 0.43-2 times the mass of the sun (without outlier)

The red points represent database mass values (Fig. 5 and 6). The red trendline represents the model function (Fig. 5 and 6). The error bars cannot be seen in the graph as they are so small they are covered by the points themselves (Fig. 5 and 6). In general, the points display a positive linear trend between M_* and the combined variable. Stars with smaller mass values (1.858×10^{30} to 4.980×10^{30} kg) lie closest to the model function. However, none of the points lies on the model function (Fig. 5 and 6). The star with the mass 3.781×10^{31} kg is an obvious outlier in the trend (Fig. 5).

8.3 Measuring the quality of the model (discrepancy)

The root mean square error value is selected to quantitatively measure the discrepancy between

the calculated mass values and the database mass values because more weight is given to larger errors (Chai & Draxler, 2014). This is important since highly accurate stellar mass values are needed to correctly predict a star's evolution and other properties such as stellar gravity (Torres *et al.*, 2009). Below is the formula for root mean square error, where E_{rms} represents the root mean square error, $M_{predicted}$ is the calculated mass values using the colour-mass models, $M_{database}$ is the database mass values, and n is the number of stars in the sample:

$$E_{rms} = \sqrt{\frac{\sum_{n=1}^n (M_{predicted} - M_{database})^2}{n}} \quad (56)$$

8.3.1 For the colour-mass relationship describing stars 2-55 times the mass of the sun

The root mean squared error of the calculated mass values using the colour-mass relationship describing stars 2-55 times the mass of the sun is 5.395×10^{30} kg, which is around 2.7 solar masses.

8.3.2 For the colour-mass relationship describing stars 0.43-2 times the mass of the sun

The root mean squared error of the calculated mass values using the colour-mass relationship describing stars 0.43-2 times the mass of the sun is 9.389×10^{30} kg or approximately 4.7 solar masses.

9 Evaluation

9.1 Evaluation of results

9.1.1 Model for stars 2-55 times the mass of the sun

As most database stellar mass values lie close to/on the model function, the model is fairly accurate (Fig. 3 and 4). The database stellar mass values increases at a rate approximately equal to the model's slope, indicating that the slope of this model ($\sqrt[3.5]{4\pi\sigma\frac{M_{\odot}^{3.5}}{L_{\odot}}b^4}$) is fairly accurate (Fig. 3

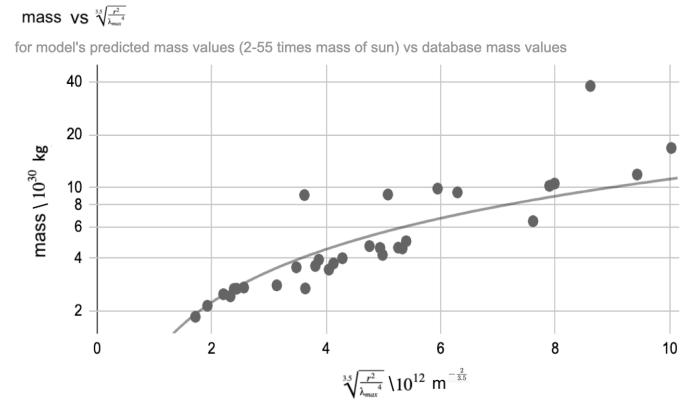


Fig. 7. Graph comparing database mass values with the model for stellar masses 2-55 times the mass of the sun using a logarithmic scale on the y-axis

and 4). Since a logarithmically scaled graph magnifies smaller discrepancies between the actual data and the model, Fig. 7 can be used to evaluate the model's accuracy. As the discrepancies between the actual data and model displayed in Fig. 7 are still relatively small, the model accurately describes the database values. However, the root mean square error of the model is 5.395×10^{30} kg, which is significantly greater than the masses of over half of the stars in the sample. Albeit the high root mean square error value, this value may be deceptive since more weight is given to larger errors, and removing the largest error would already decrease the root mean square error to 1.912×10^{30} kg, which is almost 3 times less than the original value.

9.1.2 Model for stars 0.43-2 times the mass of the sun

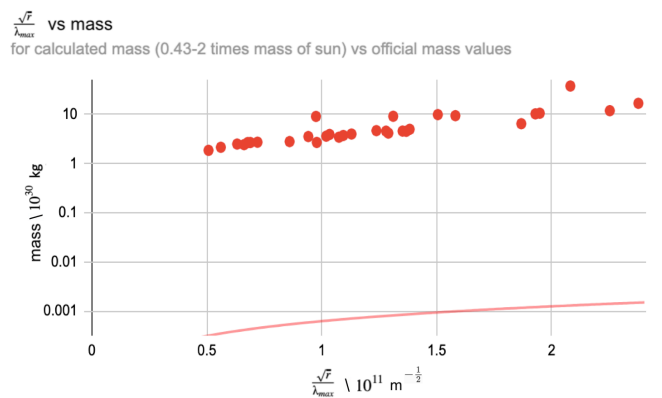


Fig. 8. Graph comparing database mass values with the model for stellar masses 0.43-2 times the mass of the sun using a logarithmic scale on the y-axis

The graph comparing database values to the model function demonstrates that the model consis-

tently underpredicts stellar mass by three orders of magnitude, indicating that the model is inaccurate (Fig. 5 and 6). From Fig. 5 and 6, it is clear that the database stellar mass values increases at a rate drastically greater than the slope of the model's function, indicating that the slope of this model

$(\sqrt[4]{4\pi\sigma\frac{M_{\odot}^4}{L_{\odot}}b^4})$ does not accurately describe the

actual data (Fig. 5 and 6). Since a logarithmically scaled graph magnifies smaller discrepancies between the actual data and the model, Fig. 8 can be used to evaluate the model's accuracy. As the discrepancies between the actual data and model displayed in Fig. 8 are large, the model does not accurately describe the actual mass values. Additionally, the root mean squared error value of the model is 9.389×10^{30} kg, which is greater than most of the stars in the sample. Albeit the high root mean square error value, this value may be deceptive since more weight is given to larger errors, and removing the two largest errors would already decrease the root mean square error to 5.931×10^{30} kg, which is slightly over half of the original value.

9.2 Evaluation of the applicability of the models

These models directly relate the colour of the star to the mass of the star and only rely on the additional parameter stellar radii, which makes them easy to apply. However, these models come with a potential caveat: prior knowledge about which mass range the star is in may be needed. This is because, during the derivation, I used mass-luminosity relationships that apply to different stellar mass classes, with both model functions producing mass values that vary by 3 orders of magnitude. Despite this, as evidenced by Fig.3, the model derived from the mass-luminosity relationship for stars 2-55 times the mass of the sun can describe stars in the 0.43-2 mass range as well. Therefore, it could be possible that the model derived from the mass-luminosity relationship for stars 2-55 times the mass of the sun is simply more accurate and thus more applicable than its counterpart.

9.3 Evaluation of the databases

Although the HYG database has a large range of data that is collected from numerous reliable databases such as the Hipparcos and Yale Bright Star Catalogue, this database also has questionable reliability as the uncertainties of the measurements are not included (Nash, 2006). If the measurement error from the measurement equipment is not known, there is a high chance that the uncertainty of the values can be underestimated. Additionally, since this database is compiled by one person, there could be errors such as wrong spectral class and Henry Draper numbers (Nash, 2006). This could cause inaccuracies in predicted trend calculations, introducing systematic error. Despite the limited sample size, Torres' database is valuable to the evaluation of the models since it is much more accurate and reliable than the HYG database as it not only used various other methods to verify the accuracy of the mass and radii values, but also displayed uncertainties in measurements, all of which were under 3% of the stellar mass/radius.

9.4 Improvements

More databases (preferably ones that state uncertainty or at least the instruments used to collect the data) could be used to cross-check the data in the HYG database to ensure the accuracy of the data. Since all the stars that were involved in this investigation were less than 20 times the mass of the sun, stars from a broader range of mass classes could be sampled to holistically evaluate the models' accuracy. The uncertainty of constants such as solar mass and luminosity which were assumed to have no uncertainty to simplify calculations, could be integrated into uncertainty calculations to prevent the underestimation of uncertainty and compute the uncertainty of the slope, allowing for a more in-depth evaluation of the models' accuracy.

10 Conclusion

Colour-mass relationships for stars based on the mass-luminosity relationships between 2-55 and 0.43-2 times the mass of the sun were successfully derived. Calculated mass values were then

compared to the database stellar mass values of 31 stars from Torres' database (Torres *et al.*, 2009). Since both the regular and logarithmically scaled graph show relatively small discrepancies between the model for stars between 2-55 times the mass of the sun and the database stellar mass values, this model is accurate (Fig. 3 and 4). However, as both the regular and logarithmically scaled graph show large discrepancies between the model for stars between 0.43-2 times the mass of the sun and the database stellar mass values, this model does not accurately describe the database stellar mass values (Fig. 5 and 6).

11 Further Enquiry

An interesting investigation could be to investigate the colour-mass relationships derived from mass-luminosity relationships that describe more extreme mass classes such as those that describe stars less than 0.43 times the mass of the sun and those that describe stars greater than 55 times the mass of the sun. This could allow for more thorough testing of the accuracy of using the colour-mass relationship method to calculate the mass of stars. A comparative study between different methods of calculating stellar mass (e.g. comparing the accuracy of using the star's velocity to derive mass with using the colour-mass relationship) would also lend insight into whether the degree of accuracy of the colour-mass relationships are acceptable or not.

References

- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7, 1247-1250. <https://www.geosci-model-dev-discuss.net/7/C473/2014/gmdd-7-C473-2014-supplement.pdf> doi: 10.5194/gmd-7-1247-2014
- Dintsios, N., Artemi, S., & Polatoglou, H. (2018). Evaluating Stars Temperature Through the B-V Index Using a Virtual Real Experiment from Distance: A Case Scenario for Secondary Education. *International Journal of Online Engineering (iJOE)*, 14, 162. doi: 10.3991/ijoe.v14i01.7842
- Duric, N. (2004). *Advanced Astrophysics*. Cambridge University Press.
- Gowdy, R. H. (2009). *PHYS103: Module 021*. <https://courses.vcu.edu/PHY-rhg/astron/html/mod/021/index3.html>
- Kepler, J. (1968). *Harmonices mundi*. Culture Et Civilisation.
- Lecchini, S. (2007). *How dwarfs became giants : the discovery of the mass-luminosity relation*. Bern Studies In The History And Philosophy Of Science.
- Nash, D. (2006). *The HYG Database | The Astronomy Nexus*. <http://www.astronexus.com/hyg>
- Russel, H. N. (1914). Relations Between the Spectra and other Characteristics of the Stars. *Nature*, 93, 227-230. doi: 10.1038/093227b0
- Salaris, M., & Cassisi, S. (2008). *Evolution of Stars and Stellar Populations*. Chichester] John Wiley Sons.
- Torres, G., Andersen, J., & Giménez, A. (2009). Accurate Masses and Radii of Normal Stars: Modern Results and Applications. *The Astronomy and Astrophysics Review*, 18, 67-126. doi: 10.1007/s00159-009-0025-1
- Tsokos, K. A. (2014). *Physics for the IB Diploma* (6th ed.). Cambridge University Press.
- Unsöld, A., & Baschek, B. (1983). *The New Cosmos : Der neue Kosmos <engl.>*. Springer.

Proving Fermat's theorem on the sum of two squares by elaborating on Zagier's "one-sentence" proof

Lok Tong Coco Yeung 楊樂同

Introduction

The Mathematics: Analysis and Approaches course at higher level places an emphasis on investigating and developing a deep understanding of mathematics (International Baccalaureate Organization, 2019), and understanding proofs is a large part of it. The proofs we learn in the classroom are challenging but straightforward: they take a while to conjecture, but the ideas are straightforward. So, when I first heard about the "one-sentence proof" by the American-German mathematician, Don Zagier, my interest was immediately piqued.

Zagier's "one-sentence" proof proves Fermat's theorem on the sum of two squares using only one sentence (Zagier, 1990). Pierre de Fermat was a 17th century mathematician who contributed greatly to the branch of Number Theory. Fermat generated many theorems in his time, such as Fermat's little theorem, Fermat's polygonal number theorem, Fermat's last theorem, *etc.*, and he would challenge his fellow mathematicians to solve them through letter. But he is also known for not revealing the proofs to his theorems. Fermat's theorem on the sum of two squares took mathematicians a century to finally discover a proof, and the fact that a mathematician can summarise the proof in just one sentence greatly amused me.

I want to take this opportunity to explore and appreciate sophisticated and concise proofs and learn about different ways to go around proving something. Pure maths and number theory fascinate me because of how complex theories can come from simple ideas. Through this exploration, I wanted to see if I can succeed in understanding complex ideas and explaining them in

simple words.

Through this investigation, I will explain Zagier's "one-sentence" proof and present all ideas as simple explanations so that the proof is more approachable for myself and others, as I believe that such brilliant ideas should be understood by non-mathematicians who may be unable to understand sophisticated mathematical terminology. In this paper, I will guide myself and others through Zagier's "one-sentence" proof, and in the end, understand the way Zagier approached this problem and appreciate the complexity of this one sentence.

1 Fermat's Theorem on the Sum of Two Squares

Every odd prime number, p , which gives a remainder of 1 when divided by 4 (*i.e.* congruent to 1 mod 4) can be expressed as:

$$p = x^2 + y^2$$

where x and y are natural numbers (Patel, 2017).

2 The First 15 Odd Prime Numbers

To start off the investigation, let us take a look at some odd primes and whether or not they can be expressed as the sum of two perfect squares.

As seen in the table above, the only odd prime numbers which can be expressed as the sum of two perfect squares are those that give a remainder of 1 when divided by 4 (for example: 5, 13, and 17). These examples demonstrate that the theorem works for the first 15 prime numbers, but does not show that it will work for all odd primes congruent to 1 mod 4, nor does it show that it will not

odd prime	mod 4	$x^2 + y^2$
3	3	N/A
5	1	$1^2 + 2^2$
7	3	N/A
11	3	N/A
13	1	$2^2 + 3^2$
17	1	$1^2 + 4^2$
19	3	N/A
23	3	N/A
29	1	$2^2 + 5^2$
31	3	N/A
37	1	$1^2 + 6^2$
41	1	$4^2 + 5^2$
43	3	N/A
47	3	N/A
53	1	$2^2 + 7^2$

Table 1. First 15 prime numbers.

work for all odd primes congruent to 3 mod 4.

3 Primes Congruent to 3 mod 4

Prime numbers that are congruent to 3 mod 4 can not ever be made up of two squares. The proof for this is simple.

In order for an odd prime number, P , to be made up of two squares, one square must be odd and the other must be even because an even number plus an odd number makes up an odd number. This means P can be re-written as

$$(2k)^2 + (2m + 1)^2$$

where k and m are positive integers.

When the expression is expanded, we get

$$4k^2 + 4m^2 + 4m + 1$$

which can be factorised to

$$4(k^2 + m^2 + m) + 1$$

which is an odd number congruent to 1 mod 4.

This proves why primes congruent to 3 mod 4 can not be expressed as the sum of two prime numbers, and why any odd sum of two squares must be congruent to 1 mod 4, but does not prove why all odd primes congruent to 1 mod 4 can be expressed this way. In order to prove Fermat's theorem, we will now move on to look at Don Zagier's "one-sentence" proof.

4 Zagier's "One-Sentence" Proof

Published in the 97th volume of The American Mathematical Monthly in February of 1990, Don Zagier's proof for Fermat's theorem on the sum of two squares states:

The involution of the finite set $S = \{(x, y, z) \in \mathbb{N}^3 : x^2 + 4yz = p\}$ define by

$$(x, y, z) \mapsto \begin{cases} (x + 2z, z, y - x - z) & \text{if } x < y - z \\ (2y - x, y, x - y + z) & \text{if } y - z < x < 2y \\ (x - 2y, x - y + z, y) & \text{if } x > 2y \end{cases}$$

has exactly one fixed point, so $|S|$ is odd and the involution define by $(x, y, z) \mapsto (x, z, y)$ also has a fixed point (Zagier, 1990).

5 Breaking Down the Proof

Initially, Zagier's proof looks intimidating because there is a lot of complex mathematical language, so I wanted to approach this proof in smaller sections. Zagier's proof has three main sections: the conditions, the piecewise function, and the conclusion.

The first line of Zagier's proof establishes the conditions. Zagier establishes the set of solutions to $p = x + 4yz$ as the finite set S . So the set includes all the possible combinations of x , y , and z for which makes the equation true. Values of x , y , and z are contained in the set of natural numbers.

The piecewise function is an involutory function, where a function is its own inverse and where applying this function twice will produce the original value. So, when set S is put through this involutory function twice, the original set is derived.

An example of an involutory function would be

$$f(x) = \frac{-1}{x}.$$

$f(x)$ is an involutory function because putting x through $f(x)$ twice gives x again, which is the original value:

$$f(f(x)) = \frac{-1}{\frac{-1}{x}} = \frac{-x}{-1} = x.$$

After defining the involution of the set S , Zagier concludes that the "fixed point", where the set derived from applying the involutory function once is the same as the original set.

6 Elaborating on the Proof

The proof, like the original theorem, assumes that p is an odd prime number, and is trying to prove that this odd prime number can be summed in two squares and is congruent to 1 mod 4. The solutions to set S all belong to the natural number series, meaning that there are no negatives or zeroes. This will be important to the later parts of this proof.

Zagier's proof centers around

$$p = x + 4yz,$$

which is an expression that does not appear in the original theorem. In the Fermat's theorem, the expression

$$p = x + y$$

is used instead. But, we know that the odd prime must be made up of one odd square and one even square (because odd plus even equals odd), so the original expression can be re-written as

$$p = x + (2y)$$

where x and y are both integers, which expands into

$$p = x + 4yy.$$

Zagier's approach was to turn $4yy$ into the more general form of $4yz$, and then prove that there can be a solution where $y = z$ because that will make a perfect square.

If $y = z$, then there must be an odd number of solutions to (x, y, z) as different sets of solutions can be paired up by swapping y and z . In this proof, we want to find the condition where there is an odd number of solutions to set S . This will occur if the involutory function can be applied to (x, y, z) once to derive the same solution.

To further elaborate, two sources were referenced. They are an essay titled "The One-Sentence Proof in Multiple Sentences" by Marcel Moosbrugger (Moosbrugger, 2020) and a blog post titled "One Sentence Proof" by user dmontealegre (dmontealegre, 2012).

6.1 The Piecewise Function

There are three parts to the piecewise function when defining the involution. When this involutory function is applied to set (x, y, z) twice, (x, y, z) will be the result. To verify this, I will go through each of the three conditions.

From this point forward, the set of solutions (x, y, z) will be denoted as set U , and the set of solutions derived when the involutory function is applied to set U once will be denoted as set U' , and the set of solutions derived when the involutory function is applied to set U twice will be denoted as set U'' .

The first condition assumes that x is smaller than y minus z . When the involutory function is applied to a set U , the result yielded is

$$U' = (x + 2z, z, y - x - z)$$

In order to put set U' through the involutory function again, we must first find the condition it satisfies. Set U' satisfies the third condition, as substituting the variables in will give $x + 2z > 2z$ which must hold true because both x and z are natural numbers. Applying the involutory function to set U' using the third condition will give

$$U'' = (x + 2z - 2z, x + 2z - z + y - x - z, z)$$

which simplifies to

$$U'' = (x, y, z)$$

which is set U . This proves that this function is an involution for all sets fitting in the first condition.

The second condition assumes that x is bigger than $y - z$ but smaller than $2y$ and

$$U' = (2y - x, y, x - y + z).$$

U' also satisfies the second condition, as $2y - x$ is bigger than $y - x + y - z$, but smaller than $2y$. Applying the involutory function again gives

$$U'' = (2y - 2y + x, y, 2y - x - y + x - y + z).$$

This simplifies to the following, which is equal to the set U .

$$U'' = (x, y, z)$$

The third condition assumes that x is bigger than $2y$ and

$$U' = (x - 2y, x - y + z, y)$$

U' satisfies the first condition because $x - 2y$ is smaller than $x - y + z - y$. Applying the involutory function to U' gives

$$U'' = (x - 2y + 2y, y, x - y + z - x + 2y - y)$$

which also simplifies to the following, which is equal to set U .

$$U'' = (x, y, z),$$

6.2 The Solution

We should remember that the point of this is to prove that there is a solution where (x, y, z) turns into (x, z, y) when passed through the involutory function once, and there should be no change to x . In the previous three paragraphs, we can actually see that if the set of solutions fit the first condition, applying the function will always make x greater, and if the set of solutions fit the third condition, applying the function will make x smaller. Thus, the only way x does not change is if the original set satisfies second conditions.

When set U satisfies the second condition, U' is $(2y - x, y, x - y + z)$. Since we want to find where x stays the same, we need to investigate where $2y - x = x$

$$2y - x = x$$

$$2y = 2x$$

$$y = x$$

This turns the equation

$$p = x + 4yz$$

into

$$p = x + 4xz,$$

which makes p divisible by x . However, since p is a prime number, and x is a natural number, the only thing x can be is p or 1. But, x can not be p because that would make the right hand side of the equation bigger than the left hand side because the right hand side must be larger than p and thus must be larger than p .

This makes

$$x = y = 1.$$

So,

$$p = 1 + 4(1)z$$

$$p - 1 = 4z$$

$$\frac{p - 1}{4} = z$$

As z is a natural number, p minus 1 divided by 4 must be a whole number. Therefore, all odd primes congruent to 1 mod 4 will have an odd set of solutions to

$$p = x^2 + 4yz.$$

Thus, all odd primes that give a remainder of 1 when divided by 4 can be the summation of two squares.

7 The Visualisation

Another way of explaining this proof is through the "windmill visualisation" (Mathologer). Let us take a look at an example: the odd prime number 17.

Here are the different solutions to the set (x, y, z) that gives $p = 17$.

$$(1, 1, 4), (1, 4, 1), (1, 2, 2), (3, 1, 2), (3, 2, 1)$$

They are solutions because

$$17 = x^2 + 4yz$$

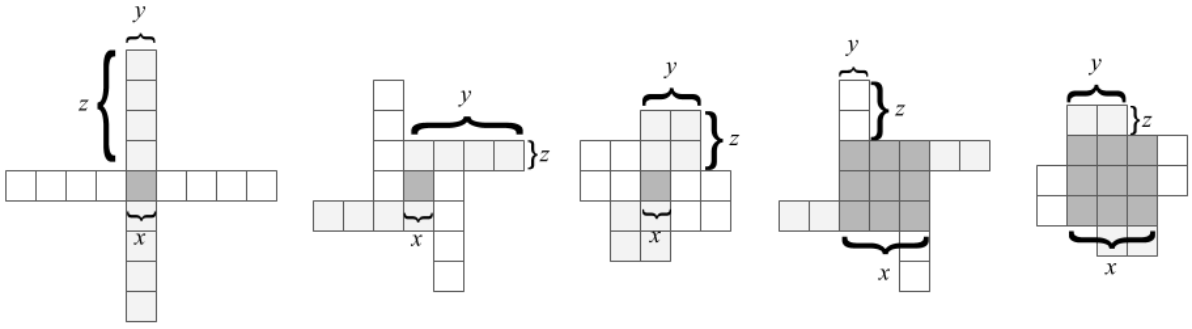


Fig. 1. Adapted from the diagrams presented in Mathologer’s Youtube Video (Mathologer, 2020).

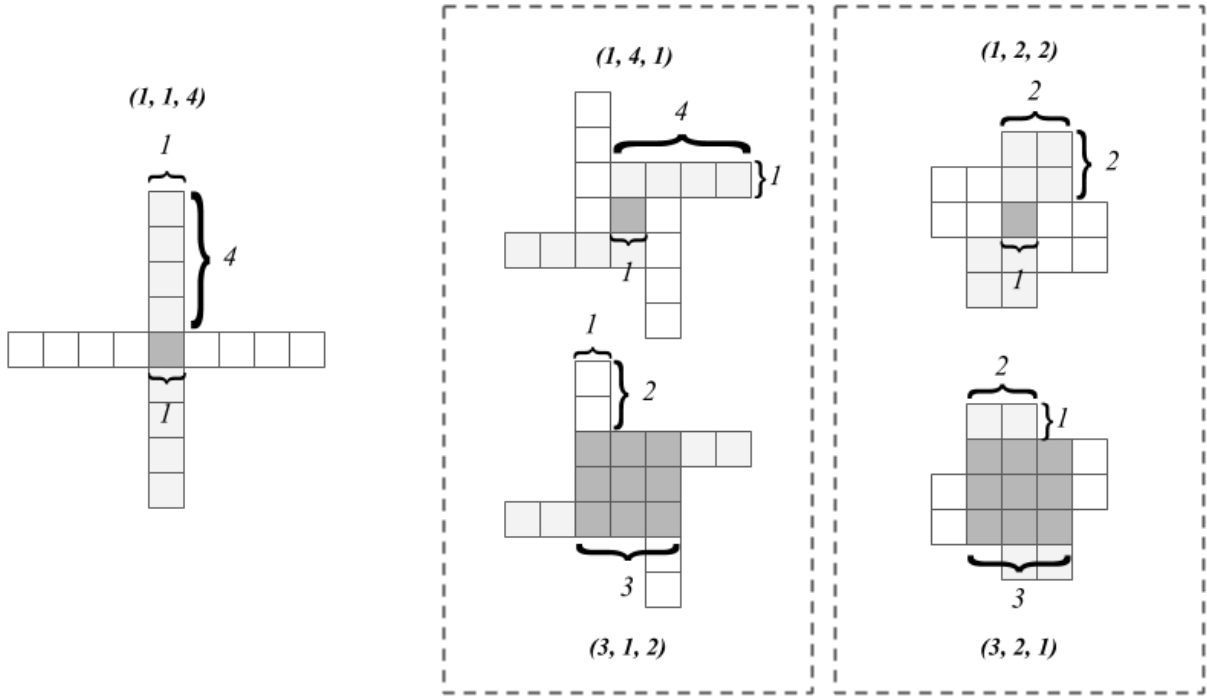


Fig. 2

and substituting the set into the equation above equates the two sides of the equation:

$$(1)^2 + 4(1)(4) = 1 + 16 = 17$$

$$(1)^2 + 4(4)(1) = 1 + 16 = 17$$

$$(1)^2 + 4(2)(2) = 1 + 16 = 17$$

$$(3)^2 + 4(1)(2) = 9 + 8 = 17$$

$$(3)^2 + 4(2)(1) = 9 + 8 = 17$$

As we can see, (1,1,4) can be paired up with (1,4,1) and (3,1,2) can be paired up with (3,2,1). The only solution that does not have a pair is (1,2,2), because in this case, y is equal to z . So, when there are an odd number of solutions, there will be one solution where y is equal to z .

This is one way of pairing up the solutions. The other way of pairing up the solutions is through the visualisation of the sets of solutions in the formation of windmills, where x is the side length of the center square, and y and z are the side lengths of the rectangles, the "blades", surrounding it.

Figure 1 depicts the five solutions in the visualised as windmills.

We can see that they can pair up where each pair has the same configuration but is a different set.

The only windmill that is unique is the one with a 1 by 1 square in the center.

Since x is the sidelength of the center square, we know that x is equal to 1. y is also equal to 1

because one of the side-lengths of the "blade" is 1. Although there can be other cross formations of windmills with bigger squares in its center, we know this is impossible because x and y must be divisible by p , but can not be larger than or equal to p . Thus, the only cross formation that exists for prime numbers is a windmill with a 1 by 1 square in its center.

Once again, this will give us

$$p = 1 + 4(1)z$$

and thus

$$\frac{p-1}{4} = z$$

This visualisation of the proof shows us how the only numbers that can have a unique square in the center must be congruent to 1 mod 4.

The piecewise function actually define pairs of windmills of the same structure but different set values. Putting the set through the involutory function once will give a new set that produces a windmill of the same shape.

For example, if we take set U to equal $(1, 4, 1)$, then it satisfies the first condition as $1 < 4 - 1$, and set U' is equal to $(3, 1, 2)$, which is shown in figure 2.

Moreover, the windmill with the 1 by 1 square in its center fits the requirements of the second condition of the piecewise function as $1 - 4 < 1 < 2$, and is the set of solutions that is its own involution.

This visual proof also more clearly illustrates how an odd prime will always produce an odd set of solutions to S , and thus having a solution where $y = z$.

Conclusion

Going into this proof, I knew it would be challenging to dissect. The mathematical terminology is sophisticated and the proof is also non-constructive, meaning that it does not give any examples, it only proves that this theorem holds true. However, once I understood the explanation

behind the proofs, I realised it was actually a pretty simple idea, and though it looks confusing at first to understand Zagier's proof, the most challenging maths you'll need to know is addition and subtraction. This greatly contrasts with other proofs of Fermat's theorem on the sum of two squares like the one derived by Euler which uses infinite descent, or Lagrange's proof which requires understanding integral quadratic forms. To me, these methods are much more challenging than dissecting what Zagier has presented in the proofs are very conceptual and difficult to visualise, unlike Zagier's which can be easily visualised as shapes.

Through this exploration, I successfully elaborated upon Zagier's "one-sentence" proof, both in words, and in graphics, which allowed me to prove Fermat's theorem on the sum of two squares.

Zagier's proof is a reminder for me to think outside the box and to think about complicated ideas in a simple way. Understanding this proof has given me a chance to extend my understanding of proofs beyond the classroom, but also to apply some of my pre-existing knowledge, including, but not limited to, understanding piecewise functions and set symbols. It has also given me a chance to delve deeper into number theory and to think of different ways to conjecture proofs, which may be beneficial to me in my future studies in this mathematics course.

References

- dmontealegre. (2012). *One sentence proof*. Word-Press. <https://danielmath.wordpress.com/2012/12/26/one-sentence-proof/>
- International Baccalaureate Organization. (2019). *International Baccalaureate Diploma Programme Subject Brief Mathematics: Analysis and Approaches*. Retrieved 2020-09-13, from <https://www.ibo.org/contentassets/5895a05412144fe890312bad52b17044/subject-brief-dp-math-analysis-and-approaches-en.pdf>
- Mathologer. (2020). *Why was this visual proof missed for 400 years? (Fermat's two square theorem)*. Retrieved 2020-09-13, from <https://www.youtube.com/watch?v=DjI1NICfjOk>
- Moosbrugger, M. (2020). *The One-Sentence Proof in Multiple Sentences*. Retrieved 2020-09-12, from <https://medium.com/cantors-paradise/the-one-sentence-proof-in-multiple-sentences-ab2657efc576>
- Patel, P. (2017). *Quadratic Reciprocity*. The University of Chicago. Retrieved 2020-09-13, from <http://math.uchicago.edu/~may/REU2017/REUPapers/Patel.pdf>
- Zagier, D. (1990). A One-Sentence Proof That Every Prime $p \equiv 1 \pmod{4}$ Is a Sum of Two Squares. *The American Mathematical Monthly*, 97, 144. Retrieved 2020-09-13, from <https://people.mpim-bonn.mpg.de/zagier/files/doi/10.2307/2323918/fulltext.pdf> doi: 10.1080/00029890.1990.11995565

How effective is Online Learning during the COVID-19 pandemic, according to student's perceptions?

Andrew Minghan Jiang 蔣明翰

Abstract

This paper investigates student perceptions and attitudes towards e-learning during the initial COVID-19 lockdown at the Independent Schools Foundation (ISF) Academy, a private secondary school in Hong Kong. A survey was administered to 96 secondary students in Apr - May of 2020, to assess the different aspects of e-learning (Self-efficacy, Academic achievement, Interaction, Motivation, Autonomy, Development of skills, Workload, Technical systems, Teaching methods, and Improvements). Overall, students self-reported the aspects to be satisfactory, rating e-learning on average 3.36 out of 5; however, when compared to traditional learning, students rated e-learning an average of 2.28 out of 5. It was found that Interaction and Motivation were most affected by the unfamiliar online medium.

Introduction

E-learning (also online learning) originated in the late 80s of the twentieth century in the United States, known as Computer-Based Training (Hubackova, 2015, p. 1189). By 2010, over 21 million US college students were enrolled in degree-granting courses (Allen & Seaman, 2013).

The World Health Organization (WHO) declared COVID-19 a global emergency on January 30th, 2020 and a global pandemic on March 11th, 2020. COVID-19 has caused school closures resulting in over 1.5 billion students staying home across 106 countries (UNESCO, 2020). In order to adapt, many schools have abruptly transitioned to deliver e-learning.

In mainland China, the Chinese Ministry of Education launched the Disrupted Classes, Undisrupted Learning initiative. Across 334 universities, surveyed students were generally satisfied with e-learning: over half found it effective. However, around 40% of students found it just satisfactory (Wu, 2020, n=256504). 36.3% of students agreed that e-learning was worse than traditional learning, and when asked separately, 28.7% of students agreed that e-learning was better than traditional

learning. Song *et al.* (2020) investigated learning motivation and perseverance in Chinese secondary students across 8 provinces (n=1117) and found that students lacked concrete learning aims, resilience, and self-discipline. A major concern for secondary students was that they lacked teacher guidance and peer interaction online (Xue & Shen, 2020, n=821). While teachers generally agreed that e-learning could improve students' independent learning abilities, they also acknowledged that it required students to be very disciplined and teachers could not manage students very effectively; three quarters of parents felt pressured to supervise their children because they lacked self-discipline.

In Hong Kong, all kindergartens, primary and secondary schools were closed from February to May 2020, July to September 2020, and November to February 2021. Lingnan University (2020) found that only 27% of Lingnan University students were satisfied with their online learning experience, and 60% of students found online learning less beneficial than classroom teaching (n=1,227). Students' largest challenges were lack of self-discipline, a poor learning atmosphere, and eye fatigue due to long screen time. In addition, Chinese University of Hong Kong (2020) asked secondary

students to rate the statement ‘online learning is valuable’ and respondents rated it an average of 2.97 out of 5 (n=1,168), which is only moderately effective. Self-reported stress and anxiety levels were 4.05 and 4.15 out of 5 respectively. These results indicate that students are still adapting to the sudden change to e-learning, and more research has to be done to make learning online more effective.

This study aims to investigate the following research questions:

1. How do ISF students feel about the skills developed through e-learning compared with skills developed through traditional learning?
2. How interactive is e-learning at ISF?

ISF This study was conducted at ISF Academy, a nonprofit IB secondary school in the Southern district of Hong Kong. ISF is a relatively new school, founded in 2003 by Nobel laureate Charles Kao, and had 1868 students as of 2018 (The Independent Schools Foundation Academy, 2018b). ISF’s class of 2019 has reported a 100% graduation rate with a mean IB score of 37 out of 45 (The Independent Schools Foundation Academy, 2018a), compared to a world average of 29.65 (International Baccalaureate Organization, 2019). Classes are taught with a mix of online and traditional tools. However, the school had not implemented any form of online learning, and classes were conducted completely in classrooms until February 24, 2020, then transitioning completely online because of the pandemic.

1 Method

A pilot study was conducted on grade 10 ISF students with a survey of seven close-ended and two open-ended questions. Feedback from students and the school prompted the investigation of the entire secondary school.

A survey of 21 questions measuring student perceptions of e-learning was designed. Aspects that we decided were worthwhile investigating included self-efficacy, a student’s belief in their capacity to reach a certain level of performance (Ban-

dura, 1997; Kauffman, 2015); perceived academic achievement, a self-reported reflection of how effectively students utilized the knowledge they learned (Alghazo, 2010); interaction, the interpersonal exchange of information, both learner-learner and teacher-learner (Mahanta & Ahmed, 2012); motivation, the internal drive and discipline required to direct and maintain effective learning (Jansen et al., 2016); autonomy, a student’s responsibility to make decisions regarding their learning (Dickinson, 1995; Jansen et al., 2016); workload, the amount of time and effort spent on school-related tasks (Muilenburg & Berge, 2005); technical systems, the infrastructure in place to support e-learning (Muilenburg & Berge, 2005); teaching methods, the strategies, approaches, and procedures used to teach Mazoue (2013); and improvements.

Survey questions were adapted from previously-validated e-learning scales Cole et al. (2014); Muilenburg & Berge (2005). Three types of questions were devised to measure the scales, general questions, comparison to traditional learning questions, and interaction-specific questions. General questions holistically measured each aspect of e-learning; respondents rated each sub-question on a Likert scale from 1 to 5 with 1 being ‘Strongly Disagree’ and 5 being ‘Strongly Agree’. Comparison questions compared respondent’s e-learning experience to their traditional one; respondents rated questions on a Likert scale from 1 to 5, 1 being ‘Online learning is much more effective’ and 5 being ‘Traditional learning is much more effective’.

To measure students’ perceptions of e-learning, Table 1 shows the ten aspects included in the survey, and examples of each. Tables 2-4 show the breakdown of more specific questions used in the survey. A table of all survey questions can be found in Appendix 1.

Indicator	Question
SE1	I expect to do well in my classes I'm confident I can excel at my online assignments and tests I do not procrastinate tasks assigned through online learning
IN1	I tend to initiate interaction with my teachers My interaction with my peers are helpful to my learning Online learning supports interaction between students and teachers
MO1	My work environment provides no distractions from online work I can remain concentrated until I am finished with online work I find online work interesting to do
AU1	I find online learning enjoyable I feel in control of my own learning I feel in control over my learning environment
DS1	I discuss ideas and knowledge with peers and teachers I initiate and further discussion I consider ideas from multiple perspectives
DS2	I am maintaining regular exercise I am maintaining sufficient sleep I am maintaining healthy self-esteem
TS1	Online learning is flexible Meeting online is convenient The school is equipped with the resources to support online learning

Table 1. General questions. A full list of the sub-questions can be found in Appendix 1.

Indicator	Question
SE2	Comparison of self-efficacy to traditional learning
AA1	Comparison of learning skills to traditional learning
IN2	Comparison of interaction to traditional learning
MO2	Comparison of motivation to traditional learning
AU2	Comparison of autonomy to traditional learning
DS3	Comparison of the development of skills to traditional learning
DS4	Comparison of online and traditional notes
WL2	Comparison of workload to traditional learning
TM1	Comparison of teaching methods to traditional learning

Table 2. Comparison to traditional learning questions

Indicator	Question
AA2	Subject-specific comparison of scores to traditional learning
IN5	Subject-specific comparison of interaction to traditional learning

Table 3. Subject-specific questions

Indicator	Question
IN3	Is asynchronous or synchronous learning more conducive for interaction?
IN4	Please rank which form of interaction you use the most

Table 4. Interaction-specific questions

2 Results

The survey was sent via school email to all students in ISF Secondary and responses were accepted from May 15 to June 3, 2020; valid responses were received from 96 students. The mean age of respondents was 13.88 years (SD: 0.090); 69.79% of respondents were female and 26.04% were male. On average, students spent 8.63 hours online per day (SD: 0.105), more than a third of each day. 79.17% of respondents spoke English at home, 54.27% Cantonese, a local dialect, and 43.75% Mandarin; 4.17% spoke another language at home (percentages do not add up to 100 because some respondents speak multiple languages at home). 33.33% of respondents only spoke 1 language at home, whereas 52.08% spoke 2, and 14.58% spoke 3.

2.1 Improvements and concerns

The following aspect was an open-ended question asking students The majority of improvements addressed classroom issues like decreasing the workload, improving lesson organization, and increasing classroom engagement. The next-most brought up were personal issues like decreasing screen time, increasing peer interaction, as well as better time-management. The final, least significant portion was technological issues, including technical issues like fixing wifi or learning-related websites, and having cameras on during lessons.

Very few respondents wanted learning to be all online (4.35%) or all traditional (17.39%). Almost

Scale	Weighted average of General & Comparison questions	Weighted average SD
SE	2.974	1.113
AA	2.373	1.090
IN	2.817	1.069
MO	2.143	1.244
AU	2.399	1.144
DS	2.924	1.223
WL	2.532	1.125
TS	3.596	0.967
TM	2.111	1.168

Table 5. Scoring of general and comparison questions. The complete results with descriptive analysis can be found in Appendix 2.

half (46.74%) of students preferred that the majority of learning should be in classrooms, with some online learning to supplement it. However, most students (74.36%) believed that online learning also had its merits.

3 Conclusion

COVID-19 has forced students to stay at home, resulting in the popularization of online education worldwide. Students generally believed online learning to be effective, rating general statements an average of 3.362 out of 5; however, when compared to traditional learning, students rated online learning an average of 2.278 out of 5.

All measured aspects of learning were less effective online, but online learning enjoyment and motivation were most decreased compared to traditional learning. Students found communication and social skills to be the most diminished when online. This may be due to a decrease in interaction when learning online. Interaction is strongly associated with online learning effectiveness, and active, regular interaction with peers and teachers can improve learning effectiveness (Kauffman, 2015; Mahanta & Ahmed, 2012; Muilenburg & Berge, 2005). The reported decrease in e-learning could be attributed to a decrease in interaction, as evidenced by the low scores of ‘Communication’ and ‘Social’ in IN2. Furthermore, most

students believed that synchronous interactions were of higher quality than asynchronous interactions. Therefore, it can be concluded that decreased interaction is one of the reasons why students believe e-learning is less effective than traditional learning; it can also be proposed that interaction greatly impacts learning motivation and enjoyment, as these aspects were the ones most affected.

Students also found it more difficult to stay on task and that their home learning environments were distracting. They also reported that it was more difficult to retain and understand information online and that discussions and in-class activities were less effective than traditional learning. This could be attributed to the online medium, as students are not in the presence of a teacher or peers and may feel physically disconnected. Physical interaction is also non-existent, which could hinder the learning process (Mahanta & Ahmed, 2012). Limited self-discipline is another concern, as not everyone can stay on-task for extended periods of time (Mahanta & Ahmed, 2012).

On the other hand, students found the interaction with their peers to be very helpful to their learning. This is reflected in the responses to ‘When I can’t understand the material, I ask another student for help’ and ‘My interaction with my peers are helpful to my learning’, both of which had very high agreement rates. Students generally agreed to the statement ‘I am keeping a network of close friends’, ‘I am satisfied with the response that I receive from my peers’, and ‘I feel cooperative toward my peers’. This indicates that although peer interaction may have decreased in quantity, it is still of high quality. Students also generally agreed that ‘Meeting online was convenient’ and that they felt in control ‘of [their] own learning’ as well as ‘over [their] learning environment’. These could both be perks of the online medium as students can learn in their home environment, which is both convenient and gives them autonomy over their learning environment.

Students also agreed to the statements ‘I expect to perform well in my classes’, ‘I do not want to waste my tuition’, ‘I feel a sense of achieve-

ment when I complete online work’, and ‘I would feel guilty not doing tasks assigned through online learning’. From a motivation perspective, three of the four above statements are intrinsic motivators, with the only extrinsic motivator was tuition. Motivation is important to online learning success as it relates to self-discipline, which is key in an environment where most of the learning is done unsupervised (Kauffman, 2015; Muilenburg & Berge, 2005; Turley & Graham, 2019). Turley & Graham (2019) also writes that a self-motivated student may learn more than what is required and overcome situations without adequate support. If students were given inspiration to become more intrinsically motivated, perhaps they might be more driven to finish assignments and achieve higher. This could be achieved by making tasks more interesting, enjoyable, or satisfying to complete (Ryan & Deci, 2000).

4 Discussion

This study investigates how effective e-learning was at a private school in Hong Kong during the first wave of COVID-19 in early 2020. Contrary to much past e-learning research, schools were not prepared for the sudden transition online because of COVID. Interaction was largely reduced because of the shift online, and although students found e-learning to be effective, when compared to traditional learning, e-learning scored worse in the aspects measured.

When compared to existing literature on e-learning, the conclusions of this study may stand out (Alghazo, 2010; Muilenburg & Berge, 2005). This may be due to COVID and the fact that schools and students were not prepared for the transition to e-learning. When compared to studies published after 2019 (Chinese University of Hong Kong, 2020; Wu, 2020; Yang *et al.*, 2020), the conclusions are much more similar. The COVID context could have a magnifying effect on the strengths and weaknesses of e-learning, and it reflects how prepared schools are to implement e-learning.

Limitations

There are several limitations of this study that should be taken into consideration in future research. First is the limited sample size of one secondary school in Hong Kong. This limits the generalizability of data collected to schools similar to this one. Second is the use of convenience sampling, which could affect internal validity, inaccurately representing the perceptions of the student body. Third is the use of a self-report survey as the only form of data collection. This limits the internal validity of the data collected, as participants may conform to demand characteristics. Fourth and finally, this study focused too much on superficial measures of e-learning (academic achievement, development of skills, workload) instead of deeper-level measures (interaction, self-discipline, academic achievement). This could affect the internal validity of results, as many aspects are interlinked, and represent e-learning effectiveness unrealistically.

Future Research

More focus on student motivation and autonomy, as well as research into aspects like self-discipline and would paint a more accurate picture of students’ perceptions of e-learning effectiveness. Furthermore, using interviews to follow-up the survey may produce more insightful, in-depth results and students can voice their thoughts on improvements to e-learning. A larger sample size would also result in a more accurate representation of students.

Appendix 1 All Survey Questions

Follow the QR code for the full survey.



Appendix 2 Survey Data

Follow the QR code for all the survey data.



References

- Alghazo, A. (2010, 01). Comparing Effectiveness Of On-line and Traditional Teaching Using Students' Final Grades. *Online Journal for Workforce Education and Development*.
- Allen, I. E., & Seaman, J. (2013, 01). *Changing Course: Ten Years of Tracking Online Education in the United States*. <https://www.bayviewanalytics.com/reports/changingcourse.pdf>. Babson Survey Research Group.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Chinese University of Hong Kong. (2020, 05). *CUHK Study Reveals that Local Secondary School Students Face Great Challenges and Pressure in Online Learning Under the Epidemic* | CUHK Communications and Public Relations Office. <https://www.cpr.cuhk.edu.hk/en/press/cuhk-study-reveals-that-local-secondary-school-students-face-great-challenges-and-pressure-in-online-learning-under-the-epidemic/>. Chinese University of Hong Kong.
- Cole, M. T., Shelley, D. J., & Swartz, L. B. (2014, 10). Online instruction, e-learning, and student satisfaction: A three year study. *The International Review of Research in Open and Distributed Learning*, 15. doi: 10.19173/irrodl.v15i6.1748
- Dickinson, L. (1995, 05). Autonomy and Motivation a Literature Review. *System*, 23, 165-174. Retrieved from <https://www.sciencedirect.com/science/article/pii/0346251X95000055> doi: 10.1016/0346-251x(95)00005-5
- Hubackova, S. (2015, 06). History and Perspectives of Elearning. *Procedia - Social and Behavioral Sciences*, 191, 1187-1190. doi: 10.1016/j.sbspro.2015.04.594
- International Baccalaureate Organization. (2019). *The IB Diploma Programme Final Statistical Bulletin*. <https://www.ibo.org/contentassets/bc850970f4e54b87828f83c7976a4db6/dp-statistical-bulletin-may-2019.pdf>. International Baccalaureate Organization.
- Jansen, R. S., van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2016, 10). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education*, 29, 6-27. doi: 10.1007/s12528-016-9125-x
- Kauffman, H. (2015, 08). A review of predictive factors of student success in and satisfaction with online learning. *Research in Learning Technology*, 23. doi: 10.3402/rlt.v23.26507
- Lingnan University. (2020, 06). *LU study reveals over 60% of Hong Kong university students have found online learning not as effective as face-to-face teaching during COVID-19*

- pandemic*. <https://www.ln.edu.hk/news/20200624/lu-study-reveals-over-60-of-hong-kong-university-students-have-found-online-learning-not-as-effective-as-face-to-face-teaching-during-covid-19-pandemic>. Lingnan University.
- Mahanta, D., & Ahmed, M. (2012). E-Learning Objectives, Methodologies, Tools and Its Limitation. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2, 46-51.
- Mazoue, J. (2013, 01). *The MOOC Model: Challenging Traditional Education*. <https://er.educause.edu/articles/2013/1/the-mooc-model-challenging-traditional-education>. EDUCAUSE.
- Muilenburg, L. Y., & Berge, Z. L. (2005, 01). Student barriers to online learning: A factor analytic study. *Distance Education*, 26, 29-48. doi: 10.1080/01587910500081269
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78. doi: 10.1037/0003-066x.55.1.68
- Song, N., Zhao, Q., & Luo, S. (2020). Status quo, Problems and Countermeasures of Middle School Students' Grit under Major Epidemic. *Research in Educational Development*, 8.
- The Independent Schools Foundation Academy. (2018a). *Impressive IB Results for ISF Students - The ISF Academy*. <https://academy.isf.edu.hk/en/2019/07/impressive-ib-results-for-isf-students/#:%7E:text=ISF%20has%20reported%20a%20100>. The Independent Schools Foundation Academy.
- The Independent Schools Foundation Academy. (2018b). *ISF School Profile 2018-19*. <https://academy.isf.edu.hk/en/category/school-publications/>. The Independent Schools Foundation Academy.
- Turley, C., & Graham, C. (2019). Interaction, Student Satisfaction, and Teacher Time Investment in Online High School Courses. *Journal of Online Learning Research*, 5, 169-198.
- UNESCO. (2020, 03). *COVID-19 Educational Disruption and Response*. <https://en.unesco.org/covid19/educationresponse>. UNESCO.
- Wu, D. (2020, 04). 疫情期間大學生線上學習調查報告 [*College e-learning report during the pandemic*]. 廈門大學教師發展2中心 [Center for Teaching and Learning Development, Xiamen University].
- Xue, L., & Shen, Q. (2020, 06). 疫情期間中學生居家學習狀況調查與分析 [*Investigation and Analysis of Secondary School Home Learning During the Pandemic*]. China Academic Journal Electronic Publishing House.
- Yang, J., Pei, W., Liu, S., Zhang, D., Jiang, L., & Yu, R. (2020, 03). 疫情時期在線教與學實踐案例與經驗 [online teaching and learning experience during the pandemic]. *中國電化教育 [China Educational Technology]*, 400, 1-13. Retrieved 2020-10-03, from <https://kns.cnki.net/kcms/detail/11.3792.G4.20200317.1105.004.html>

Modelling COVID-19 using the SIHD model

Ying Chun Justin Man 文言中

1 Introduction

Witnessing how COVID-19 has been spreading across the world, it has piqued my interest in simulating the mathematical model of such a fatal pandemic. After reading a lecture by HKU that studies the Hong Kong 2003 SARS outbreak with the standard SIR model (Susceptible, Infectives and Recovered), I was intrigued by the model's simplicity and practicality to approximate the pandemic trend. By replacing the category "Recovery" and adding new categories – Hospitalised and Death, as well as introducing the Basic Reproductive Number, this investigation enhances the design of SIR model based on the real-life context of coronavirus. Up-to-date statistics will also be referenced in the chosen parameters to ensure the model is a good replication, if not, a rough approximation of real-life trends. To conclude, a holistic judgement on this model's significance and veracity (e.g. limitations and flaws in assumption) will be discussed.

2 SIR Model

2.1 Background

The standard SIR model is a compartmental model which consists of three deterministic differential equations, to compute the dynamics of SIR in a closed population over time (Wikipedia Contributors, 2019). It is useful in estimating the duration, spreading rate, epidemiological parameters (e.g reproductive number) of the pandemic and how government health interventions can affect its outcome (Ng, 2004).

2.2 Assumptions

The standard SIR model is built based on the following assumptions:

1. **Homogenous mixing:** Each susceptible individual faces an equal risk in getting infected.
2. **Population size remains constant:** It assumes that the epidemic is sufficiently short so that any change in population due to births, death, immigrants or emigrants are negligible. This is also to reduce excessive variables and keep the model as simplified as possible.
3. **Constant rate of recovery:** Similar to the one above, it also assumes that infectives "recover" at a constant rate and people can only be infected once. (i.e. they can no longer be infected, they have either recovered and developed immunity, or died)

2.3 Variables

Since the model is estimating the pandemic trend over a period of time, we will set time (measured in days) as the independent variable t . Hence, the dependent variables SIR can be expressed the following below:

S_t – The number of individuals that are susceptible/capable of getting the disease, but not yet infected

I_t – The number of individuals that are currently infected by the disease

R_t – The number of individuals that have recovered, or died due to the disease

Based on the assumption of constant population size, we can express these variables in relation to the total population, N , such that the sum of the total number of susceptible, infectives and recovered is equal to the total population.

$$S_t + I_t + R_t = N$$

2.4 Developing Equations

By assuming homogeneous mixing that each individual faces an equal risk in getting the infection, the interaction between the susceptible and infected people can be expressed with the two variables' product: $S_t I_t$. We can then now develop a system of differential equations of the standard SIR model, which represents the rate of change of each variable, governs the behaviour of the pandemic over a period of time. (i.e. the rate of change of these variables over time)

The rate of change of the number of susceptible people, as shown in equation 1, is negative because the total number of susceptible will reduce over time. The variable α represents the probability for susceptible people getting infected, which can be influenced by the frequency of interaction with infected people.

$$\frac{ds}{dt} = -\alpha SI \quad (1)$$

Consequently, the rate of change in infectives over time, as shown in equation 2, will equal to that proportion of susceptible transferred to the infected people. At the same time, some infected individuals can also recover at a rate γ , meaning that there is a loss in infected people and they are assumed to remain immune to future infection.

$$\frac{dI}{dt} = -\alpha SI - \gamma I \quad (2)$$

Hence, this brings us to the last equation in equation 3, in which the rate of change in recovered individuals over time will be a gain in the infected people at the given rate γ .

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

3 Improving the SIR model based on the actual COVID-19

Although the standard SIR model can act as a simple tool to estimate the crude trend of an epidemic, it needs to be recognized that the model does not accurately represent the actual picture of coronavirus as it is over-simplified. For example, the

coronavirus has a relatively high death rate (approximately 4%), compared to the flu and common cold (the usual ones in SIR model case studies). Also, the assumption of recovered people have perfect immunity is invalid. A recent study by Melbourne university found out that the median immunity of a recovered patients to the virus is only 14%, meaning that there is still a high chance of getting re-infected (Juno *et al.*, 2020).

3.1 Three Distinct Changes

1. I_t is subdivided into “infected” and “hospitalised” – which will be labelled I and H respectively. This is because the recovery rate can differ among the two groups of infected people. According to the WHO, 80% of the infected people can recover from the disease without needing special treatment (World Health Organisation Indonesia, 2020). And it typically takes them 2-3 weeks to recover, whilst those with severe infection are around 6 weeks.
2. **The transferal of recovered to susceptible** – as mentioned above, there is an equal chance of recovered patients to get infected again. Hence, recovered people will eventually return to the susceptible category.
3. **Death category** – is added in this model. It assumes patients that have serious illness will seek hospitalization ($I \rightarrow H$), and they will either recover and move back to the susceptible category, or to the death category.

3.2 Defining the New Variables

γ_1 – the recovery rate to transfer from I to S

γ_2 – the recovery rate to transfer from H to S

β – the rate of hospitalisation from I to H

δ – the death rate from H

The flow of SIHD can be expressed using a flow diagram:

We can now develop a new system of differential equations:

$$\frac{ds}{dt} = -\alpha SI + \gamma_1 I + \gamma_2 H$$

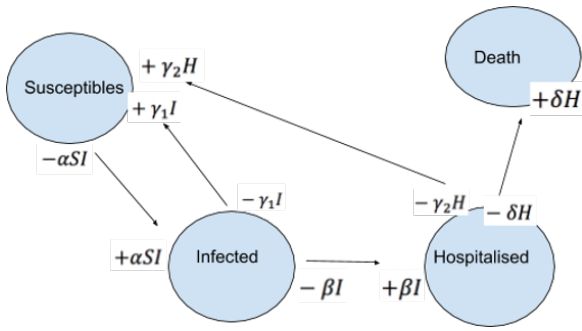


Fig. 1. A flow diagram of SIHD model

$$\frac{dI}{dt} = \alpha SI - \gamma_1 I - \beta I$$

$$\frac{dH}{dt} = \beta SI - \gamma_2 I - \sigma H$$

$$\frac{dD}{dt} = \delta H$$

4 Solving the Differential Equations Using Euler's method

In reality, these differential equations are often difficult and complex to solve. Hence, the Euler's method, an iterative approximation tool that allows us to solve these first-order ordinary differential equations, is chosen in this investigation to allow us graph the dynamics of our SIRD model. First of all, the general form of Euler's differential equation for y and t can be expressed as below:

$$\frac{dy}{dx} = f(t, y)$$

$$y(t_0) = y_0$$

The equation $f(t, y)$ is a known function, and the initial conditions are also known values. Using these formulae as our starting point, we can generate the rest of the solution by using iterative formulas, where Δx is the step size between each intervals. The smaller the step size, the more accurate the approximation is.

$$x_{n+1} = x_n + \Delta x$$

$$y_{n+1} = y_n + \Delta x \frac{dx}{dy}$$

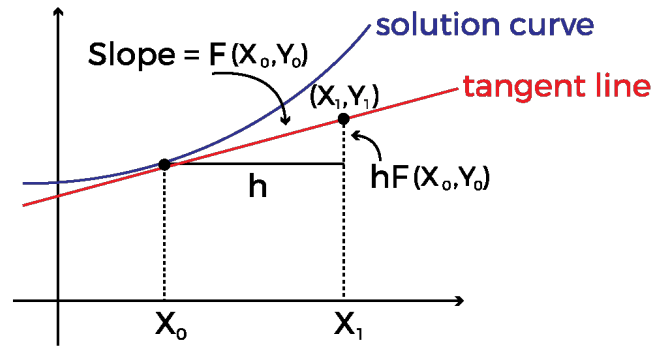


Fig. 2. Euler's Method, taken from (CalcWorkshop, 2019).

This method can also be thought of using tangent lines over a given step size Δx (it is expressed as h in figure 2) to approximate the solution of a point, given its initial value (CalcWorkshop, 2019). Hence, applying this method to our SIHD model, we will first have these equations:

$$S_t = S_{t-1} + \Delta S$$

$$I_t = I_{t-1} + \Delta I$$

$$H_t = H_{t-1} + \Delta H$$

$$D_t = D_{t-1} + \Delta D$$

By applying the same method, we can also find the next point solution:

$$S_{t+1} = S_t + \Delta S$$

And then we add our variables in the SIRD model.

$$S_{t+1} = S_t - \alpha S_t I_t \times \Delta t + \gamma_1 I_t \times \Delta t + \gamma_2 H_t \Delta t$$

Taking out the common factor:

$$S_{t+1} = S_t - (\alpha S_t I_t - \gamma_1 I_t - \gamma_2 I_t) \times \Delta H$$

Repeating the same for $I_t H_t D_t$.

$$I_{t+1} = I_t + I_t (\alpha S_t - \gamma_1 - \beta) \times \Delta t$$

With these four equations, we can choose some values for our parameters, and then we can construct a table to show/calculate the incremental changes in SIHD with each step size. In this investigation, I will use the US as an example because it has a huge population size and they currently have the most COVID-19 cases worldwide.

4.1 The Rationale Behind the Chosen Parameters

1. $S_0 = 3.28 \times 10^8$
 - (a) US entire population without vaccine/immunity
2. $I_0 = 53$ cases
 - (a) Number recorded on the 24th of February, it is the approximate time when local community transmission of coronavirus started (Worldometer, 2020b).
3. $H_0 = 24$ cases
4. $D_0 = 0$ cases
5. $\alpha = \frac{1}{10} = 0.1$
 - (a) The typical incubation time of Coronavirus symptoms range from 5-14 days (Lauer *et al.*, 2020), hence taking the average $\frac{(5+14)}{2} = 9.5$ days. This is rounded it up to 10 days, so that calculations are easier.
6. $\gamma_1 = \frac{1}{14} \approx 0.0714$ (correct to 3 sig.fig.)
 - (a) The median time from onset to clinical recovery for mild cases is approximately 2 weeks (World Health Organisation, 2020).
7. $\gamma_2 = \frac{1}{31.5} \approx 0.0318$ (correct to 3 sig.fig.)
 - (a) Three to six weeks for patients with severe or critical disease, hence taking the average $\frac{(3+6)}{2} = 4.5$ weeks = 31.5 days.
8. $\beta = 0.1$
 - (a) This value depends on the frequency of testing, so infected people (who have severe symptoms) can be transferred into the hospitals

- (b) The US currently has less than 10% of daily positive testing (Johns Hopkins University Coronavirus Resource Center, 2020).

9. $\delta = 0.03$

- (a) Depends on this formula $\frac{\text{daily death rate}}{\text{daily hospitalised cases}}$
- (b) The value fluctuates each day, so I took some values from each month and found the average.
- (c) For example:
 - i. US May 24th: $\frac{627}{19934} \approx 0.0315$
 - ii. March 24th: $\frac{269}{11254} \approx 0.0239$
- (d) $\Delta t = 1$
 - i. Measured in days, step size is one day.
 - ii. Step size 1 day is chosen because cases are usually reported daily.

4.2 Approximating Covid-19 Cases in the US Using Euler's Method

Note: values for $S I H D$ are expressed in $x \times 10^8$.

When $t = 0$

$$S_t = 3.28$$

$$I_t = 5.3 \times 10^{-7}$$

$$H_t = 2.4 \times 10^{-7}$$

$$D_t = 0$$

When $t = 1$

$$S_1 = S_0 - (0.1 \times S_0 \times I_0 - 0.0714 \times I_0 - 0.0318 \times H_0) \times 1$$

$$I_1 = I_0 - I_0(0.1 \times S_0 - 0.0714 - 0.1) \times 1$$

$$H_1 = H_0 + (0.1 \times I_0 - 0.03175 \times H_0 - 0.03 \times H_0) \times 1$$

$$D_1 = D_0 + 0.03 \times H_0 \times 1$$

When $t = n$

$$S_n = S_{n-1} - (0.1S_{n-1}I_{n-1} - 0.0714I_{n-1} - 0.0318H_{n-1})\Delta n$$

$$I_n = I_{n-1} - I_{n-1}(0.1S_{n-1} - 0.0714 - 0.1)\Delta n$$

$$H_n = H_{n-1} + (0.1I_{n-1} - 0.03175H_{n-1} - 0.03H_{n-1})\Delta n$$

$$D_n = D_{n-1} + 0.03H_{n-1}\Delta n$$

By computing these equations and parameter values into an Excel spreadsheet, we can create a graph showing the dynamic of SIHD curves, as shown in figure 3.

With our chosen parameters, figure 3 displays the trend and changes in Susceptible, Infected, Hospitalized and Dead people in the US over a period of time. Looking at the I_t curve, a rapid increase can be observed starting from around day 89, and it eventually hits its maximum at day 125, with 71.5 million people infected. Simultaneously, H_t is growing at an exponential rate, peaking at day 139, with 86.9 million people hospitalized.

Upon first glance, the graph doesn't make too much sense if it is compared to the real-life statistics. This model predicts that there will be an exponential growth in D_t , and eventually starting to flatten out at day 293, where the death cases will account for almost half of the population. However, if we think about it, the epidemic will be disastrous if we have half of the US population wiped out at the end! One reason is because we assume that the parameters will remain unchanged over these 240 days, which is unlikely given that the government has been implementing strict health regulations to limit the rate of interaction and parameter α . To present this on a graph, I reduced the parameter α from 0.1 to 0.05 after day 50 (to explain this conceptually: the government noticed a surge in cases, so they implemented health regulations after 50 and half of the interactions were cut off to reduce the number of cases).

Now we can now construct another graph based on that, as shown in figure 4.

Figure 4 shows the effect of reducing the parameter α . From day 1-50, when the parameter α is 0.1, a clear exponential growth in I_t can still be seen, perhaps due to the lack of health regulations and public awareness of the virus. And at day 50, the curve suggests that there is a total of 38752 new cases, which is fairly reasonable, considering the number of cases in the US in real life (around 30000 cases per day in April). After the implementation of health regulations, a kink from day 50 to 51 and a decline in new cases the next day can be observed. This is probably because people are forced to stay home, wear masks and use hand sanitizers. After day 50, there is a steady decline in new daily cases, and the number of hospitalisation is also decreasing after day 82. However, it is also worth mentioning about the flaw of this graph - it is unlikely in real life that these regulations will already have such immediate impact on reducing the number of cases the next day it is implemented. Nevertheless, this still showcases the powerful impact of reducing the parameter α (*i.e.* the need for government regulations).

On the other hand, for the D_t curve, it is still increasing because the epidemic is still spreading at a relatively fast rate. To understand the reason why, we can first look at the basic reproductive number.

5 The Basic Reproductive Number

Another way to look at how the epidemic will spread is through the Basic Reproductive Number, which refers to the number of new infections caused by an individual. This can be estimated by the expression (Jones, 2007):

$$\begin{aligned} & (\text{average infection caused by} \\ & \text{an individual per time unit}) \\ & \times (\text{expected duration of infection}) \end{aligned}$$

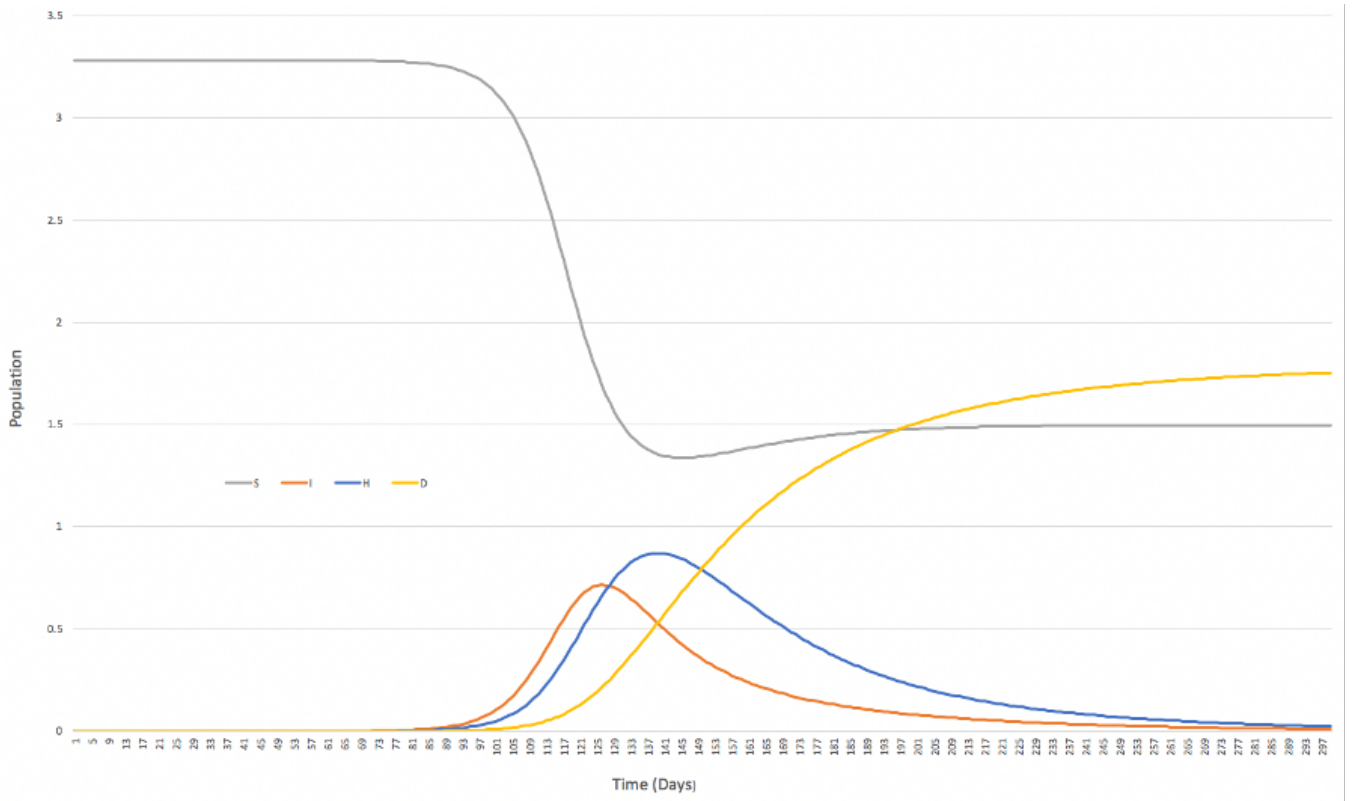


Fig. 3. Proposed SIHD model of the US pandemic outbreak. To provide a more holistic overview, the x -axis interval is automatically set to 4 days by Excel.

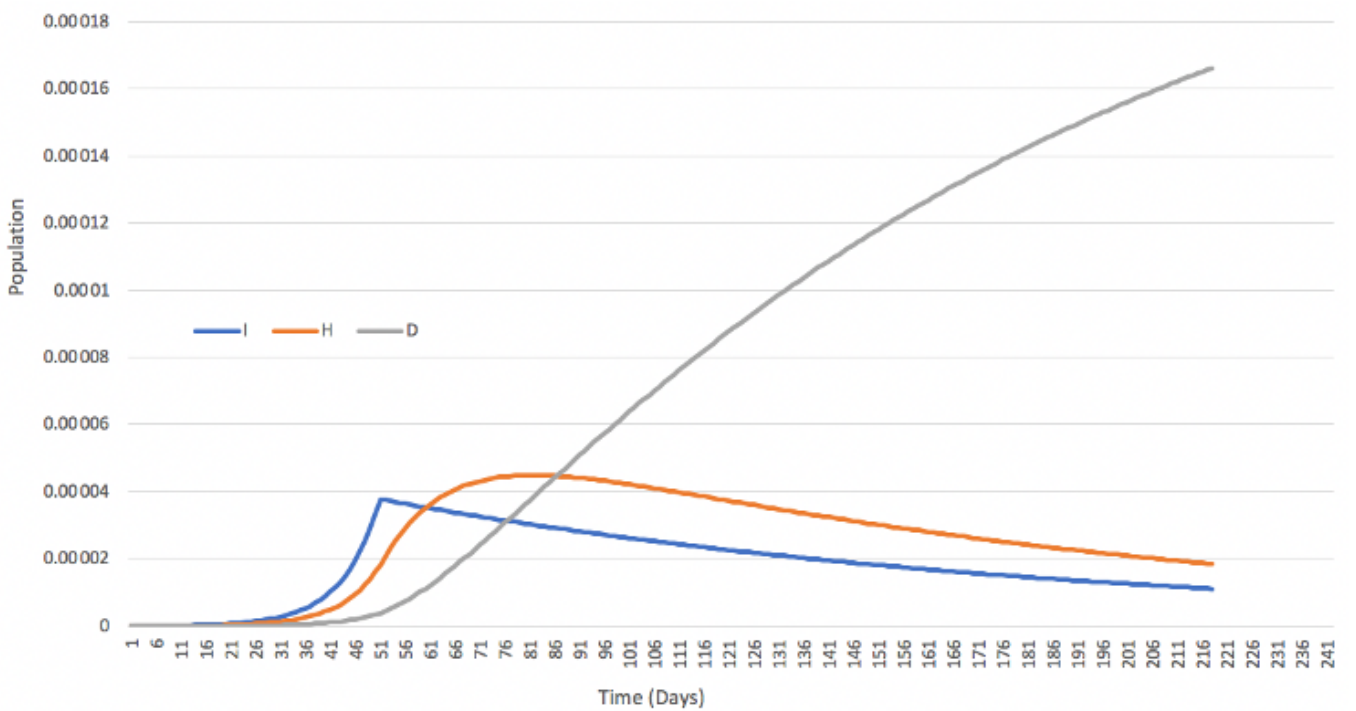


Fig. 4. Implementing health regulations after 50 days. S_t is not shown because its size is too large relative to IHD.

5.1 Finding the Expected duration of Infection

To find the expected time taken for infection, recall that each infected individual has a probability of γ_1 to recover in each time step (section 3.2).

This allows us to construct the following tree diagram, as shown in figure 5.

If we define the discrete random variable X be the number of time steps it take a patient to recover, then we can construct the following probability ta-

ble for X .

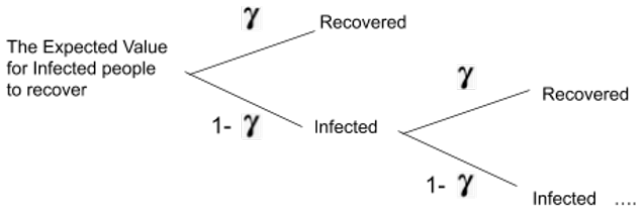


Fig. 5. Tree diagram

X	1	2	3	n
$P[X = x]$	γ_1	$\gamma_1(1 - \gamma_1)$	$\gamma_1(1 - \gamma_1)^2$	$\gamma_1(1 - \gamma_1)^{n-1}$

Table 1. Probability table for X .

Hence, we can calculate when the sum of n tends to infinity.

$$\begin{aligned}
 E[X] &= \sum_{x=1}^{\infty} x \times P[X = x] \\
 &= \gamma_1 + 2(1 - \gamma_1)\gamma_1 + 3(1 - \gamma_1)^2\gamma_1 + \dots \\
 &\quad \dots + n(1 - \gamma_1)^{n-1}
 \end{aligned}$$

This sequence is neither arithmetic nor geometric. But it can be solved by separating it into sections:

$$= \gamma_1 + (1 - \gamma_1)\gamma_1 + (1 - \gamma_1)^2\gamma_1 + (1 - \gamma_1)^3\gamma_1 \dots \quad (1)$$

$$+ (1 - \gamma_1)\gamma_1 + (1 - \gamma_1)^2\gamma_1 + (1 - \gamma_1)^3\gamma_1 \dots \quad (2)$$

$$+ (1 - \gamma_1)^2\gamma_1 + (1 - \gamma_1)^3\gamma_1 \dots \quad (3)$$

⋮

Expression 1 is now a geometric series, with first term γ_1 and common ratio $1 - \gamma_1$, thus the sum to infinity formula can be applied:

$$S_{\infty} = \frac{a_1}{(1 - r)}$$

$$\frac{\gamma - 1}{(1 - (1 - \gamma_1))} = 1$$

Meanwhile, expression 2 has first term $(1 - \gamma_1)\gamma_1$ and common ratio $1 - \gamma_1$.

$$\frac{\gamma_1(1 - \gamma_1)}{(1 - (1 - \gamma_1))} = 1 - \gamma_1$$

This extends likewise to expression 3.

$$\frac{\gamma_1(1 - \gamma_1)^2}{1 - (1 - \gamma_1)} = (1 - \gamma_1)^2$$

Now looking at the product of each expression, we can notice another infinite geometric sequence, thus we can apply the sum to infinity formula again.

$$\begin{aligned}
 E[X] &= 1 + (1 - \gamma_1) + (1 - \gamma_1)^2 \dots + (1 - \gamma_1)^n + \dots \\
 &= \frac{1}{1 - (1 - \gamma_1)} = \frac{1}{\gamma_1}
 \end{aligned}$$

As a result, the expected duration of each infected person to recover should be $\frac{1}{\gamma_1}$.

5.2 Finding the Average Infection Caused by the Individual Per Time

For the average infection caused by the individual per time, this can be calculated as follows. Since there is no vaccine available currently, we can assume that the initial condition where $t = 0$, is $S_0 = N$ (the total population) because everyone can be susceptible to the disease. We can also deduce that $S_0 > S$ because the number susceptible will decrease over time. Hence, taking the equation $\frac{dI_1}{dt}$, we can substitute the initial condition S_0 and rearrange:

$$\frac{dI_1}{dt} < I_1(\alpha S_0 - \gamma_1) - \beta I_2$$

The term $-\beta I_2$ is negligible in this context, because we assume that hospitalised infected people will be separated and have minimal or no contact with the susceptible people, thus they will not have an impact on the spread of the epidemic. As a result, we have this inequality.

$$\frac{dI_1}{dt} < I_1(\alpha S_0 - \gamma_1)$$

Focusing on the variables α and γ_1 , then rearranging them, we obtain the following.

$$\alpha S_0 - \gamma_1 > 0$$

$$S_0 > \frac{\gamma_1}{\alpha}$$

To further simplify the inequality, we multiply both sides by α and then divide by them γ_1 .

$$R_0 = \frac{S_0 \alpha}{\gamma_1} > 1$$

The ratio $\frac{S_0 \alpha}{\gamma_1}$ is known as the basic reproductive number R_0 . The value of R_0 helps us understand what is happening with the epidemic (Gog, Thomas, & Freiberger, 2020).

1. If $R_0 > 1$, it means that the epidemic will grow.
2. If $R_0 = 1$, it means that the epidemic is plateauing.
3. If $R_0 < 1$, it means that the epidemic will decline.

To prevent mass infection, the government has several ways to intervene and implement policies so that $R_0 < 1$. Given that the R_0 will be smaller when γ_1 increases, the government can decrease the value of R_0 by facilitating the development of better medical treatments, *e.g.* medicine, so that the expected duration for infected people recovery rate will decrease.

However, in the short run, the government should prioritise reducing the probability for susceptible people to get infected, *i.e.* parameter α , by implementing social distancing policy, or encourage the use of hand sanitizers. A long term solution is to reduce the number of susceptible, S_0 , by using vaccines so that more people are immune to the virus.

5.3 Application of R_0

To provide a better picture of how R_0 works, we can substitute some numbers into the equation.

Substituting our initial parameters in figure 3 gives us the equations below.

$$R_0 = \frac{S_0 \alpha}{\gamma_1} > 1$$

$$R_0 = \frac{3.28 \times 0.1}{0.0714} \approx 4.60 > 1$$

With the result larger than 1, this indicates that the epidemic will grow. Given that it is difficult to reduce S_0 (*e.g.* vaccines) and γ_1 (*e.g.* better medicine) in the short run, our only short term solution is to reduce parameter α , so that R_0 is < 1 .

$$R_0 = \frac{3.28 \alpha}{0.0714} < 1$$

Rearranging this equation, we obtain the following.

$$\alpha < 0.0218$$

Therefore, in order for the epidemic to stop spreading, we need to have parameter α to be smaller than 0.0218, which means that it needs to be roughly $\frac{1}{5}$ of the initial parameter α (0.1). Hence, going back to figure 4, this explains why the D_t is still increasing at a fairly steady rate, simply because the regulations are not strict enough, and the contact rate between susceptible and infected is still fairly high. Therefore, if we reduce the parameter to 0.0218, the graph should look like this.

Figure 6 shows a sharp decline in I_t after day 50, when α has been reduced to 0.0218. As expected, the epidemic starts to die out, with the infected and hospitalised cases reaching 0 at approximately day 121 and 145. Simultaneously, D_t also starts to flatten out at day 127, meaning that there are barely any deaths. In fact, if we look at some real-life examples, a similar pattern can be observed in China.

After the coronavirus outbreak in January, China imposed strict, or even extreme, lockdown rules on cities. For example, in Wuhan, the government closed all transport links and banned citizens from making any non-essential trips out of apartments (Culver & Westcott, 2020). Despite the lockdown's economic cost (Hubei's economy shrank by 40%), it was arguably successful as the country was able to reduce the number of cases under a short period of time (see Fig. 7). Therefore, this example emphasizes the importance of reducing the parameter α , and it is possibly one the US can learn from.

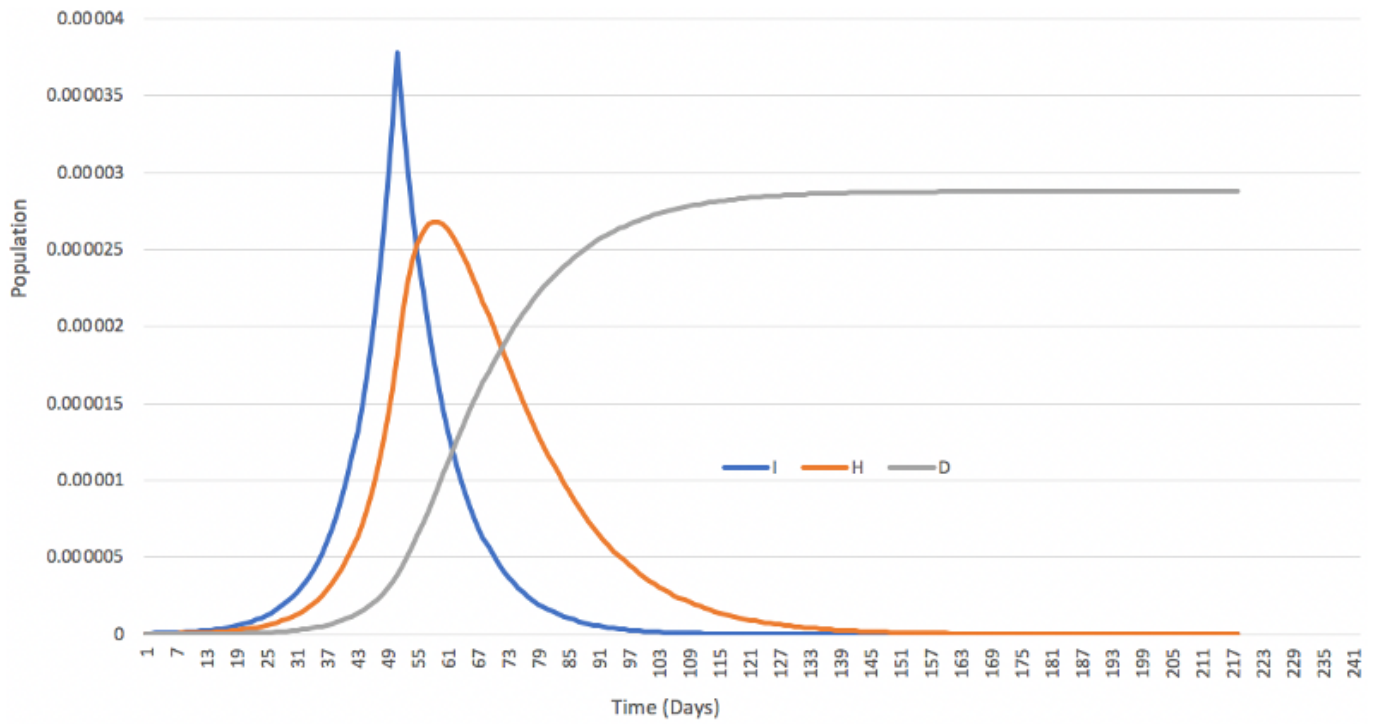


Fig. 6. Impact of strict lockdown (parameter α reduced to 0.0218)

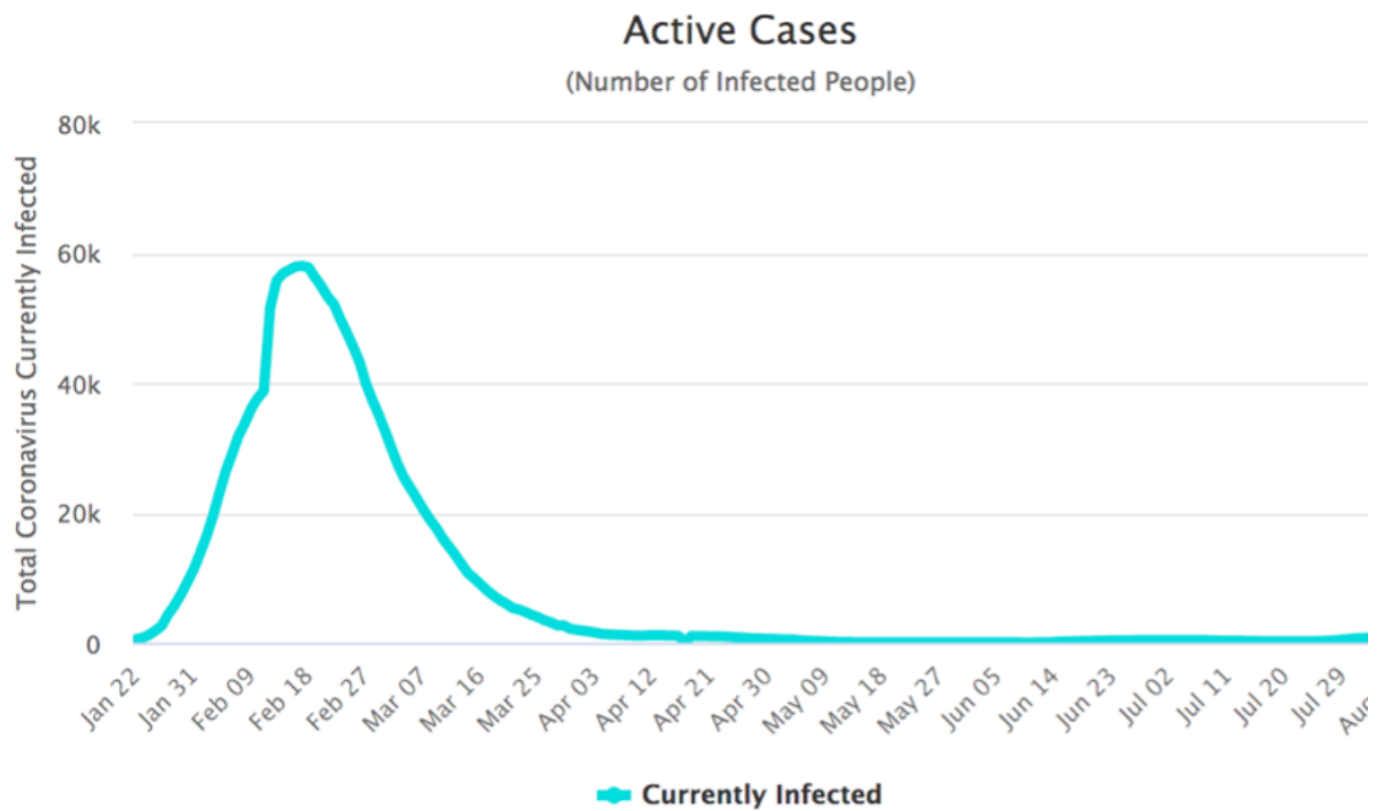


Fig. 7. Infected Cases in China from January to July (Worldometer, 2020a).

6 Conclusion

Overall, my investigation using the SIHD model and basic reproductive number is successful in approximating the current situation of the coronavirus. Even though the chosen parameter values might not be so accurate, it does help to demonstrate the significance of government regulations to reduce contact rate (parameter α), especially when vaccines and medicines will not be available in the short run.

However, this doesn't mean that the model is without its flaws. There are three major flaws in the assumptions: Firstly, it assumes that the population is homogeneously mixed, such that individuals have an equal probability of interacting with each another. This does not hold true in reality because not all susceptible individuals will have contact with other people. Instead, it is likely that the spreading rate will slow down in the long run, as most people usually live within a fixed community so there should be less mobility between different groups. Secondly, my SIHD model assumes that hospitalized people will have no further impact on the number of infected cases, because medical staff should be well protected with medical gear while hospitalized patients are insulated from the public. However, this assumption has proven to be wrong recently, as an increasing number of medical staff have become infected in Hong Kong. This flaw can simply be improved in my existing model by adding more variables/categories. Third, the assumption of constant population size (since any immigration or emigration is negligible) is also inaccurate, because the coronavirus has already become a global pandemic – any immigration will add uncertainty in the spread of virus. If foreign infected people are allowed to enter other countries, it will likely result in another wave of local community outbreak. For example, Hong Kong's third wave of coronavirus outbreak in August was due to loopholes in the city's border measures. Hence, this leads us to the major limitation of the model: in the case of coronavirus, without vaccines, it is difficult to predict the duration of the pandemic. This model cannot predict when the different waves of community outbreaks will occur, especially when

scientists are concerned that infected cases will soon surge again in winter (Roberts, 2020).

Despite the model's limitations and the unpredictable nature of pandemic, one has to understand that it is virtually impossible to predict the future behaviour of coronavirus with perfect certainty or accuracy. This is because SIR model is deterministic in nature, and it greatly relies on existing data, as well as the assumption of *ceteris paribus*, with most things being equal. In reality, getting infected should not be directly proportional, and does not follow a specific infection/spreading rate. It should in fact be a random probability, with the magnitude of it depending on various factors (such as age). Therefore, one possible improvement is a change to a probabilistic model, followed by binomial distribution and several simulations to find out the average, worst and best case scenario. However, this does not necessarily mean that the veracity of model should be underestimated. Ultimately, it still allows some flexibility for adjustment in the model (*e.g.* adding variables, restricting the size of parameter at a given time), as well as its ability to compare different policies effectiveness. And these predictions do play a role for politicians to make a policy decision, and which aspect of lockdown policies need work.

Nevertheless, through this investigation, I realised how real-life situations and mathematics can be closely connected. I was able to apply my knowledge (the Euler's Method, differential equations and geometric distribution) and improve upon the standard SIR model, while fulfilling my interest in exploring the impact of lockdown policies on the coronavirus, all of which enhanced my analytical skills and ignited my passion to further explore other aspects of the pandemic.

References

- CalcWorkshop. (2019). *How to do Euler's Method? (Simply Explained in 4 Powerful Examples)*. Author. Retrieved 2020-09-05, from <https://calcworkshop.com/first-order-differential-equations/eulers-method-table/>
- Culver, D., & Westcott, B. (2020). *Wuhan is on a slow path back to normality after 76-day coronavirus lockdown*. CNN. Retrieved 2020-09-05, from <https://edition.cnn.com/2020/04/23/asia/wuhan-coronavirus-after-lockdown-intl-hnk/index.html>
- Gog, J., Thomas, R., & Freiberger, M. (2020). *The growth rate of COVID-19*. University of Cambridge. Retrieved 2020-09-05, from <https://plus.maths.org/content/epidemic-growth-rate>
- Johns Hopkins University Coronavirus Resource Center. (2020). *How Does Testing in the U.S. Compare to Other Countries?* Johns Hopkins University Medicine. Retrieved 2020-09-05, from <https://origin-coronavirus.jhu.edu/testing/international-comparison>
- Jones, J. H. (2007). *Notes On R_0* .
- Juno, J. A., Tan, H.-X., Lee, W. S., Reynaldi, A., Kelly, H. G., Wragg, K., ... Wheatley, A. K. (2020). Humoral and circulating follicular helper T cell responses in recovered patients with COVID-19. *Nature Medicine*, 26, 1428–1434. <https://www.nature.com/articles/s41591-020-0995-0> doi: 10.1038/s41591-020-0995-0
- Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., ... Lessler, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, 172. <https://annals.org/aim/fullarticle/2762808/incubation-period-coronavirus-disease-2019-covid-19-from-publicly-reported> doi: 10.7326/M20-0504
- Ng, T. W. (2004). *The mathematics of diseases - On Modeling Hong Kong's SARS Outbreak*. Retrieved 2020-09-05, from <https://hkumath.hku.hk/~ntw/SCNC1001-2004b.pdf>
- Roberts, M. (2020). Winter wave of coronavirus 'could be worse than first'. *BBC News*. Retrieved 2020-09-05, from <https://www.bbc.com/news/health-53392148>
- Wikipedia Contributors. (2019). *Compartmental models in epidemiology*. Wikimedia Foundation. Retrieved 2020-12-05, from https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology
- Wikipedia Contributors. (2020). *Geometric distribution*. Wikimedia Foundation. Retrieved 2020-09-05, from https://en.wikipedia.org/wiki/Geometric_distribution
- World Health Organisation. (2020). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. World Health Organisation. Retrieved 2020-09-05, from <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>
- World Health Organisation Indonesia. (2020). *Media Statement: Knowing the risks for COVID-19*. World Health Organisation. Retrieved 2020-09-05, from <https://www.who.int/indonesia/news/detail/08-03-2020-knowing-the-risk-for-covid-19>
- Worldometer. (2020a). *China Coronavirus Cases*. Dadax Limited. Retrieved 2020-09-05, from <https://www.worldometers.info/coronavirus/country/china>
- Worldometer. (2020b). *United States Coronavirus*. Dadax Limited. Retrieved 2020-09-05, from <https://www.worldometers.info/coronavirus/country/us/>

A Note about Style

Articles included in this publication are written for many different purposes. Any differences in style are due to the need to adhere to the format required for that purpose. Generally, the Modern Language Association (MLA) citation and format style (8th Ed.) is used for articles written in English as part of the Oxford University Shuyuan Classics Summer Program or the NRI Scholar's Retreat (Needham Research Institute, at Cambridge University), while the STEM articles adhere to the American Psychological Association (APA) citation and format style (7th Ed.). Articles written in Chinese use footnotes following the style outlined in the Bulletin of the Institute of Chinese Literature and Philosophy. However, articles that were originally submitted as partial fulfillment of the International Baccalaureate (IB) programmes, such as the Middle Years Programme's (MYP) Personal Project or the Diploma Programme's (DP) Extended Essay, have followed the specific requirements as outlined by the student's supervisor, and they are published in this journal as they were originally submitted.

關於文體的說明

本出版物中的文章是為許多不同目的而寫的。任何風格上的差異都是由於需要遵守該目的所需的格式。一般來說，牛津大學書院經典暑期班或劍橋大學 NRI 研究所（Needham Research Institute）暑期班的英文文章，採用現代語言協會（MLA）的引文和格式（第 8 版），而 STEM 文章則採用美國心理學會（APA）的引文和格式（第 7 版）。用中文撰寫的文章採用中研院《中國文哲研究集刊》的腳注樣式。但是，如果是作為國際文憑課程（IB）的部分內容而提交的文章，如中學課程（MYP）的個人項目或文憑課程（DP）的擴展論文，則按照學生導師提出的具體要求，按原樣在本刊發表。

Needham Research Institute (NRI)

李約瑟研究所

The following article was written as a culminating essay for the Shuyuan NRI Scholar's Retreat, 2020.

以下文章是在 2020 年舉行的李約瑟研究所暑期研究項目中所寫的結題論文。

- Smallpox and its associated political implications in Qing Dynasty China (William Tristan Lee III 李欣隆)

American Geophysical Union

美國地球物理學會

The research for these articles were prepared for a poster presentation at the AGU (American Geophysical Union) Fall Meeting, 2020.

這篇文章的研究是為 2020 年 AGU (美國地球物理學會) 秋季會議上的海報演講準備的。

- Hybrid Sentiment Analysis system to extract Language Bias in news media (Sally Sijie Song 宋思婕)
- How effective is Online Learning during the COVID-19 pandemic, according to student's perceptions? (Andrew Minghan Jiang 蔣明翰)

Extended Essays 拓展論文

These articles were written in partial fulfillment of the IB Extended Essay for Grade 12 students, May 2020, or May 2021.

以下文章是 2020 年 5 月，或 2021 年 5 月畢業的 12 年級學生寫的拓展論文 (Extended Essay)。

- The antimicrobial effects of Ethanol (Chrystal Li 李瑩瑩)
- A data-driven study on the effects of changes in local sea surface temperature on the health of Coral Reefs in Hoi Ha Wan, Hong Kong (Julie Tam 譚幸臨)

Internal Assessments 內部評估論文

These articles were written in partial fulfillment of the IB Mathematics Internal Assessment for Grade 12 students, May 2021. 這是 2021 年 5 月畢業的 12 年級學生寫的 IB 數學內部評估論文。

- Fourier Series modelling with applications to Electrocardiogram readings (Dionne Daiyin Yeung 楊岱殷)

- Proving Fermat's theorem on the sum of two squares by elaborating on Zagier's "One-Sentence" proof (Lok Tong Coco Yeung 楊樂同)
- Modelling COVID-19 using the SIHD model (Ying Chun Justin Man 言中)

This article was written in partial fulfillment of the IB Physics Internal Assessment for Grade 12 students, May 2021.

這是 2021 年 5 月畢業的 12 年級學生寫的 IB 物理內部評估論文。

- A study on the relationship between the Colour and Mass of selected Main Sequence Stars (Candace Yan Yue Chung 鍾欣瑜)

This article was written in partial fulfillment of the IB Design Technology Internal Assessment for Grade 12 students, May 2021.

這是 2021 年 5 月畢業的 12 年級學生寫的 IB 設計內部評估論文。

- Design Technology: EpiPen (Stanley Ip 葉宇軒)

Theory of Knowledge Essays IB 知識論

These articles were written in partial fulfillment of the IB Theory of Knowledge Essay for Grade 12 students, May 2020, or May 2021.

這是 2020 年 5 月，和 2021 年 5 月畢業的 12 年級學生寫的 IB 知識論 (Theory of Knowledge) 論文。

- "Statistics conceal as much as they reveal." (Denton Philtjens 費丹成)
- Within Areas of Knowledge, how can we differentiate between Change and Progress? (Candace Yan Yue Chung 鍾欣瑜)

The Independent Schools Foundation Academy
1 Kong Sin Wan Road, Pokfulam, Hong Kong
Tel +852 2202 2000
Fax +852 2202 2099
Email enquiry@isf.edu.hk

