Chapman University

Chapman University Digital Commons

Computational and Data Sciences (MS) Theses

Dissertations and Theses

Fall 12-2021

An Information-Theoretic Analysis of Adherence to Physical Exercise Routines

Lily Foster

Chapman University, foste155@mail.chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/cads_theses

Part of the [Data Science Commons](#)

Recommended Citation

L. Foster, "An information-theoretic analysis of adherence to physical exercise routines," M. S. thesis, Chapman University, Orange, CA, 2021. [https://doi.org/ 10.36837/chapman.000332](https://doi.org/10.36837/chapman.000332)

This Thesis is brought to you for free and open access by the Dissertations and Theses at Chapman University Digital Commons. It has been accepted for inclusion in Computational and Data Sciences (MS) Theses by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

An Information-Theoretic Analysis of Adherence to Physical Exercise Routines

A Thesis by

Lily Clare Daphne Foster

Chapman University

Orange, CA

Schmid College of Science and Technology

Submitted in partial fulfillment of the requirements for the degree of
Master of Science in Computational and Data Sciences

December 2021

Committee in charge:

Erik Linstead, Ph.D., Chair

Vincent Berardi, Ph.D.

Rene German, M.S.

CHAPMAN UNIVERSITY
SCHMID COLLEGE OF SCIENCE AND TECHNOLOGY

Computational and Data Sciences

The thesis of Lily Clare Daphne Foster is approved.

Digitally signed by Erik
Linstead
Date: 2021.12.10
14:53:17 -08'00'

Erik Linstead

Erik Linstead, Ph.D., Chair

Vincent Berardi, Ph.D.

Rene German, M.S.

December 2021

An Information-Theoretic Analysis of Adherence to Physical Exercise Routines

Copyright © 2021

by Lily Clare Daphne Foster

ACKNOWLEDGEMENTS

I would like to thank the researchers at Arizona State University and all those involved with the WalkIT study for allowing me the opportunity to analyze your study's data, along with Dr. Vincent Berardi for his guidance and mentorship throughout this research.

Thank you to my friends in the Machine Learning and Affiliated Technologies (MLAT) lab for taking me under your wing and giving me a place on campus to call home. I'd especially like to thank my advisor and mentor Dr. Erik Linstead for believing in me and seeing my potential. There really are no words to express the impact you've had on my life.

I would like to also thank Abigail Atchison for sparking my love of computer science and showing me that it's never too late to start learning something new.

To my family and friends, thank you for your unwavering love and support throughout this process.

To my friends in the CADS program, I cannot thank you enough for everything you've taught me these past few years and especially for teaching me to believe in myself.

It truly does take a village, and I am incredibly grateful for mine.

ABSTRACT

An Information-Theoretic Analysis of Adherence to Physical Exercise Routines

by Lily Clare Daphne Foster

One of the most common recommendations in healthcare is to simply form healthy habits, but little research has been done to understand the formation and continuation of a healthy habit that isn't heavily influenced by an individual's interpretation. Arizona State University's WalkIT study aimed to analyze how goal setting and financial reinforcement can influence moderate-to-vigorous physical activity (MVPA) in adults, while using data from accelerometers to alleviate individual bias. In this trial, 512 insufficiently active adults were recruited to wear an accelerometer for 1 year and were then randomly assigned to one of the four study groups. Each group had either a static or adaptive MVPA goal and received financial rewards either based on a predetermined steadily increasing schedule or earned points towards their financial rewards by achieving their set MVPA goals. Data was collected from each of the 512 participant's accelerometer for every minute of every day for the full duration of the 1-year study indicating their physical exertion levels. In this thesis, that data is analyzed to investigate how habitual participants are with their physical exercise routines. We hypothesize that the more consistent participants are with the time of day they exercise, the more they will exercise. However, with the use of information entropy, a linear regression model, and the R^2 statistic we disprove this hypothesis. Instead, we find that the more a participant exercises, the less habitual they are with the time of day they exercise. Given how busy our lives tend to be, this aligns with the idea that

we are more likely to exercise if we fit it in whenever we have the time to, rather than attempting to stick to exercising solely at a specific time each day that will inevitably become busy when something unexpected pops up. With the natural chaos of the society we live in today, it's easier to meet goals that have built-in flexibility rather than trying to control the uncontrollable.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENTS	IV
ABSTRACT	V
LIST OF TABLES	IX
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XI
LIST OF SYMBOLS	XII
1 INTRODUCTION	1
2 THE STUDY	3
2.1 Overview.....	3
2.2 Recruitment.....	3
2.3 Design	5
2.3.1 Static Physical Activity Goals	6
2.3.2 Adaptive Physical Activity Goals.....	6
2.3.3 Immediate Financial Reinforcement.....	6
2.3.4 Delayed Financial Reinforcement.....	7
2.3.5 Behavioral Maintenance Follow-Ups	8
3 DATA	9
3.1 Overview.....	9
3.2 Preparing for Analysis	10
4 METHODS	11
4.1 Information Entropy.....	11
4.2 R-Squared (<i>R</i>²).....	12
4.3 Linear Regression	13
5 RESULTS	14
5.1 Aggregate Values by Study Group	14
5.2 Count of Mean Entropy Values	15
5.3 Linear Regression Model Application	17
6 DISCUSSION	20

7 CONCLUSION	22
REFERENCES.....	23

LIST OF TABLES

	<u>Page</u>
Table 3-1: Each 1-minute observation grouped to create 4 bins for time of day.....	10
Table 4-1: Interpreting the Entropy Value $H(\mathbf{x})$	12
Table 5-1: Aggregate Values by Trial Group	14

LIST OF FIGURES

	<u>Page</u>
Figure 1: Example of Facebook Recruiting Ad [7]	4
Figure 2: 2x2 Trial Groups Outlined	8
Figure 3: Mean Entropy Value Count for All Participants	16
Figure 4: Mean Entropy Value Count by Trial Group.....	17
Figure 5: Linear Regression Model for all Participants.....	18
Figure 6: R2 Values Including 12-, 18-, and 24-Month Follow-Ups	19

LIST OF ABBREVIATIONS

<u>Abbreviation</u>	<u>Meaning</u>
The WHO	The World Health Organization
ASU	Arizona State University
MVPA	Moderate-to-Vigorous Physical Activity
PA	Physical Activity
RSS	Residual Sum of Squares
TSS	Total Sum of Squares
WalkIT	Walking Intervention Through Texting”

LIST OF SYMBOLS

<u>Symbol</u>	<u>Meaning</u>
R^2	Statistical score measuring the fit of data to a model
β_0	Coefficient denoting the y -intercept
β_1	Coefficient denoting the slope
x	Represents an independent variable
y	Represents a dependent variable

1 Introduction

According to the World Health Organization (WHO), 28% of the world's adults aged 18 and over are not active enough with levels continuing to rise, as discovered in a 2016 study [1]. This means that they were not meeting the global recommendation of at least 150 minutes of moderate-intensity physical activity per week, or 75 minutes of vigorous-intensity physical activity per week. For decades people from all different professions have tried to address this issue, but with very little change. Numerous studies have also been done using a variety of different designs to explore the effects of anything from motivation, positive reinforcement, financial incentives, schedules, and even environmental cues [2, 3, 4, 5]. However, most of these studies rely on participants self-reporting their physical activity, leaving the intensity of the exercise up to interpretation [6]. This bias also makes it very difficult to accurately compare participants within a study since everyone has different physical capabilities. This is what researchers at Arizona State University aim to do differently in the WalkIT study. Instead of relying on the interpretation of self-reported questionnaires, participants wear an accelerometer around their wrist to collect concrete data. The design and length of this study set it apart from any similar studies done in the past with the same aim of increasing moderate-to-vigorous physical activity (MVPA) in insufficiently active adults.

This thesis utilizes the data collected from the WalkIT study, analyzing 512 datasets for each of the participants, all containing over 525,600 observations for every minute of the 1-year intervention and over 20 features, with the aim of understanding how habitual people are around the time of day they exercise.

For every unique participant, we broke up each day into 4 bins (12am-6am, 6am-12pm, 12pm-6pm, 6pm-12am), summing all bout minutes for each bin to show the total number of MVPA minutes per time of day. With this, we then calculated the daily value for information entropy, to represent the total amount of information, or uncertainty, carried in the variable. Average information entropy and average bout minutes were the two main variables used throughout this thesis to analyze how consistent participants were with the time of day they exercise. We hypothesize that there will be a linear relationship between average entropy value and average bout minutes, signifying that the more habitual a person is with the time of day that they exercise, the more they will exercise overall. After analyzing the data for all 512 participants, the results we found showed that the more a person exercises, the less habitual they are in the time of day they exercise. In the pages that follow, we will further detail the design of the WalkIT study (chapter 2), the data used for analysis (chapter 3), the methods leveraged to create and evaluate a linear regression machine learning model (chapter 4), the results obtained (chapter 5), their significance (chapter 6), along with some final concluding thoughts (chapter 7).

2 The Study

2.1 Overview

The “Walking Intervention Through Texting” (WalkIT) Arizona study aimed to increase the level of moderate-to-vigorous physical activity (MVPA) among insufficiently active adults, digging into the effectiveness of different forms of motivation using financial incentives. In Maricopa County, Arizona, 512 insufficiently active adults were recruited to take part in the 2x2 factorial trial, with the first aspect of the trial being the goal type (static or adaptive), and the second being the type of financial reinforcement system (immediate or delayed).

2.2 Recruitment

Participants were recruited for the study mainly through advertisements on Facebook. Facebook was chosen as the main advertising channel because it allows for targeted ads, where criteria can be specified to limit the scope and better reach an advertisement’s target audience. This became extremely useful for this particular study given the specificity of the eligibility requirements. An example of one of the ads shown on Facebook is shown in Figure 1 below.

Figure 1: Example of Facebook Recruiting Ad [7]

To be eligible to participate, individuals were required to be (1) an adult between the ages of 18 and 60, (2) living in one of the pre-specified neighborhoods within Maricopa County, Arizona, (3) insufficiently active as determined by one of two questionnaires (International Physical Questionnaire [IPAQ] short form [8] and/or Stanford Usual PA Questionnaire [9]), (4) generally healthy (without any major health conditions including but not limited to cancer, heart conditions, lung conditions, exercise-induced asthma, or a history of any such conditions), (5) not currently pregnant or planning on becoming pregnant within the next 2 years, (6) not currently participating in any physical activity, diet, or weight loss programs, (7) willing to wear an accelerometer daily for 1 full year, (8) access to and knowledge of a smartphone and willing to send and receive multiple text messages daily for 1 year, (9) not planning to travel outside of the region for more than 30 consecutive days, (10) not planning to move within the next 2 years.

The Facebook ad referred potential participants to the official WalkIT Arizona website (www.walkitarizona.org) where they were asked to take a survey to determine their eligibility, as well as ensuring that only one person per household was recruited. Of those that clicked on the Facebook ad, 9,475 began the screening survey, 5,294 met all the requirements, 728 attended an initial visit, and 512 were officially admitted to the study and randomly assigned to one of the four trial groups. The study took place from May 2016 to May 2018, with participant's start dates sampled equally across calendar months.

Prior to being randomly assigned to a trial group, each participant was required to complete a baseline office visit to define their individual threshold equating moderate-to-vigorous physical activity. To find each participant's unique 'cut-point', they participated in a staged treadmill test while wearing an accelerometer on their wrist, and an indirect calorimetry device to collect breath-by-breath VO_2 data while walking/running at various levels of speed [10]. This was then followed by 10 trial days to work out any issues before officially beginning the study.

2.3 Design

For the duration of the study, participants wore a blinded accelerometer on their wrist daily to gather data on their levels of movement and exertion. Participants were then required to sync their accelerometer at least once per day, sending their data into the cloud where ASU researchers could monitor each participant's progress.

As a 2x2 factorial trial, the study contained four trial groups, each with 128 randomly assigned participants [7]. The trial tested two physical activity goal conditions (Static vs. Adaptive), and two financial reinforcement conditions (Immediate vs. Delayed) outlined below.

2.3.1 Static Physical Activity Goals

Those in the static goal group were asked to achieve at least 30 minutes of MVPA each day, for at least 5 days per week throughout the entire 1-year intervention. This goal was set based on the current guidelines of physical activity being at least 150 minutes of MVPA per week [11].

2.3.2 Adaptive Physical Activity Goals

Unlike in the static goal group, participants in the adaptive goal group were given a different goal each day of the study, depending on their individual performance over the previous 9 observations. Since the goals would change every day, each goal was only valid for the 1 day it was assigned for and was personalized for every participant in the group individually. Each goal was calculated by taking the duration of MVPA minutes for the past 9 days of observations, ranking the days from lowest to highest, and finding the 60th percentile which would become the next goal. Due to this system of calculations, an individual's new goal could be higher, lower, or the same as the day before.

2.3.3 Immediate Financial Reinforcement

Participants in the immediate financial reinforcement group had the opportunity to earn points each day of the study depending on whether they met their MVPA goal that day, and if they had successfully synced their accelerometer. Earning one point was equivalent to earning \$0.01, with points being redeemable for an e-gift card once they reached 500 points (\$5.00 e-gift card). After reaching 500 points, they were given the option of choosing an e-gift card from a catalog of 12 available retailers through Tango Inc., including retailers such as Amazon, Sephora, Home

Depot, Target, and Starbucks. However, for those in the immediate reinforcement group, the magnitude and frequency of their reinforcement varied as participants progressed through a series of stages, referred to as “reinforcement stages” [12]. To progress to the next stage, participants would need to successfully reach their physical activity goal 24 times.

To develop a strong and consistent correlation between positive reinforcement and goal attainment, stage 1 began as a fixed magnitude stage where every goal attained resulted in 100 points earned. After 24 goals were successfully met, the participant would move on to stage 2 in which reaching their physical activity goal still consistently resulted in points earned, but the number of points varied from 25 to 250 points. In stage 3 and every following stage, a goal met no longer resulted in the guarantee of points earned, with the likelihood of a participant earning points steadily decreasing with each consequent stage. With there being 365 days in the study, participants had the opportunity to earn up to \$365 if they met their goal every day. However, given the design of the reinforcement schedule in this group and based on previous studies, most participants were expected to achieve, on average, 40%-73% of their goals [13, 14].

2.3.4 Delayed Financial Reinforcement

In contrast, participants in the delayed financial reinforcement group did not receive points, instead following a set 60-day escalating financial reward schedule. The requirements for this group were to wear the accelerometer for at least 10 hours a day on 4 out of the last 7 days and syncing their device. Participants in this group were not required to meet their MVPA goals to receive their rewards. If both criteria were met, participants would receive cash-equivalent e-gift cards on an escalating magnitude basis of \$15 in month 2 (M2), \$30 in M4, \$50 in M6, \$75 in

M8, and \$95 in M10. Participants had the opportunity to earn up to \$265, equaling the maximum amount the immediate reinforcement group was expected to earn on average.

	Adaptive PA Goals	Static PA Goals
Immediate Financial Reinforcement	Group A	Group B
Delayed Financial Reinforcement	Group C	Group D

Figure 2: 2x2 Trial Groups Outlined

2.3.5 Behavioral Maintenance Follow-Ups

After the end of the 1-year intervention, participants were encouraged to continue their MVPA over the next year, with follow-ups performed at 12, 18, and 24 months. At 12 months, participants went through an additional baseline visit to remeasure their individual ‘cut-point’ to define MVPA. For participating in the baseline stage, participants received a \$20 e-gift card, followed by the opportunity to earn \$25 for the 12-month measure, \$20 for the 18-month measure, and \$25 for the 24-month measure. At each follow-up stage, participants were sent an accelerometer in the mail to be worn for up to 10 days before syncing their data and returning it again in a prepaid envelope.

3 Data

3.1 Overview

The data provided by Arizona State University contained an individual file for all 512 unique participants of the study. These files were each comprised of approximately 525,600 observations, one observation for each minute of every day, with some participants having more than the expected 365 days and some having less. For the purpose of this thesis, the relevant variables included the participant's study group, date and time, count vector magnitude (amount of activity), activity level (whether that observation qualified as MVPA), the cut point used for that particular participant (based on their baseline study, what count vector magnitude value qualifies as MVPA), bout minute (if the 3 or more consecutive observations qualified as MVPA), and wear (if the accelerometer was worn).

While the data had already been mostly cleaned before we received it, an unexpected issue we did run into was that of time zones. Normally, California and Arizona being on two different time zones would not have had an impact on the fundamental analysis being performed, and both programming languages used (R and Python) would've easily adjusted [15, 16]. However, in this particular case, California observes daylight savings time while Arizona does not. With the data containing observations for a full year, the transition between time zones resulted in one hour being mismatched, thus mismatching the observations. After uncovering the problem and understanding the reasoning behind it, we were able to simply adjust the time zones and account for the lack of daylight savings time in the original data.

3.2 Preparing for Analysis

To effectively analyze the data by time of day, the first challenge became how best to define each “time of day”. Given that we had an observation for each minute, we started by first adding each bout minute for every hour of a 24-hour period, to have 24 groups (“bins”). It quickly became obvious that comparing 24 bins was not a realistic course of analysis since people often exercise across hours, for example 5:30pm-6:30pm. We then performed trial analyses for groups varying between 2 and 12 bins, eventually finding that 4 bins made the most sense to allow for the realistic variables we all encounter in our everyday lives. The bins used are outlined in Table 3-1 below.

Bin	Time
“Early Morning”	12am – 6am
“Morning”	6am – 12pm
“Afternoon”	12pm – 6pm
“Evening”	6pm – 12am

Table 3-1: Each 1-minute observation grouped to create 4 bins for time of day.

4 Methods

4.1 Information Entropy

To investigate how habitual each participant was in the time of day that they exercised, we first needed a metric that would utilize the entirety of the extensive data collected for each participant. After applying several different methods, Shannon’s theory of “Information Entropy” became the most fitting solution. In 1948, a mathematician by the name of Claude Shannon published a paper in the Bell System Technical Journal called “The Mathematical Theory of Communication” [17], where he outlined a formula to describe the average level of “information” or “uncertainty” inherent in a variable. This formula is now known as “Shannon’s Entropy” or “Information Entropy”, as it will be referred to in this thesis. Information Entropy is calculated using the following equation,

$$H(x) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (1)$$

where $H(X)$ is the amount of information carried in an event or variable X , $\sum_{i=1}^n$ denotes the sum over the variable’s n number of possible values (x_i), and $P(x_i)$ representing the probability that x_i will occur [18] [19].

For our data, we used the sum of bout minutes in each of the four bins and applied each of the four values into the information entropy formula. If all four values were zero, meaning that the participant did not exercise at all that day, a value of “-1” was returned and that day was filtered

out. Now having one value for each day, we found the mean value to find each participant's overall average information entropy value.

Information Entropy Value	Level of Uncertainty	Level of Habituality
$H(x) = 0$	Less Uncertainty	More Habitual
$H(x) = 1$	More Uncertainty	Less Habitual

Table 4-1: Interpreting the Entropy Value $H(x)$

4.2 R-Squared (R^2)

R^2 is a statistical measure representing the proportion of variance for a dependent variable as explained by an independent variable, calculated below [20].

$$R^2 = 1 - \frac{RSS}{TSS} \quad (2)$$

The Residual Sum of Squares (RSS) is the total amount of variance in the data, which isn't already explained by a regression model.

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

The Total Sum of Squares (TSS) tells us how much variation there is in the dependent variable.

$$TSS = \sum (y_i - \bar{y})^2 \quad (4)$$

The R^2 statistic is often used to assess the fit of a model to its' data, with a greater R^2 value indicating a better fit to the model.

4.3 Linear Regression

Simple linear regression is a machine learning model that attempts to predict a liner relationship between two variables [21]. This is a commonly used technique to make predictions about additional data points by seeking a trend in the data. This relationship can be expressed when an increase in one variable corresponds with a proportional increase or decrease in the other. Simply put:

$$[Variable\ 2] = [Coefficient] \times [Variable\ 1] \quad (5)$$

Mathematically, a linear regression model can be expressed through the below equation,

$$y = \beta_0 + \beta_1 x \quad (6)$$

where x is the independent variable, y is the dependent variable, β_0 is the y -intercept, and β_1 is the slope. The R^2 statistic is a common method used to measure the fit of data to a linear regression model.

When comparing the average number of bout minutes and average information entropy value it was hypothesized that there would be a linear relationship, indicating that the more habitual a person is, the more they exercise.

5 Results

5.1 Aggregate Values by Study Group

Due to the large sample size of participants in this study, we first found the values for average entropy and average bout minutes for each individual participant to be used for analysis. Initially, we broke up each participant into their trial group and analyzed the four different groups individually. This allowed for a comparison of the different motivational methods highlighted in the outline of the original study. When finding the mean and median values for each group's average entropy and bout minutes, as shown in Table 5-1, Group A showed consistently higher values compared to the rest of the groups, with the only exception being Group C having a slightly higher median for average entropy (+0.0016).

	Average Entropy		Average Bout Minutes	
	Mean	Median	Mean	Median
Group A	0.3026	0.2911	31.7370	27.3936
Group B	0.2846	0.2836	29.7330	26.7089
Group C	0.2879	0.2927	27.6203	26.6203
Group D	0.2753	0.2732	25.8851	25.6230

Table 5-1: Aggregate Values by Trial Group

5.2 Count of Mean Entropy Values

After identifying the aggregate values per study group, we next wanted to see the mean entropy distribution for all participants. Since each participant had an overall average entropy value calculated, as described above, we then counted the number of participants whose average entropy value fell within each 0.1 bin. As shown in Figure 3, this uncovered a bell-curve, indicating the possibility of a Gaussian, or normal, distribution. However, when performing the Shapiro-Wilk statistical test to evaluate the likelihood that the data was drawn from a Gaussian distribution, we found a p-value of 0.001. With an alpha value of 0.05, we can conclude that the data is likely not a Gaussian sample. However, upon further investigation, we can see that our data is clearly skewed to the left given that groups 0.7-0.8, 0.8-0.9, and 0.9-1.0 all have a count of 0, indicating that all participants did show at least some levels of habituality in their exercise routines.

Figure 3: Mean Entropy Value Count for All Participants

By further breaking up the participants into the 4 different trial groups and again identifying the counts for each 0.1 of mean entropy value, we can still slightly see the same bell-shaped curves, as shown in Figure 4. Interestingly, groups A and B are much more evenly distributed across the first 6 bins, whereas groups C and D both have one bin with a significantly higher count compared to the other bins. Groups C and D are also the only 2 groups to any participants with a mean entropy value above 0.6.

Figure 4: Mean Entropy Value Count by Trial Group

5.3 Linear Regression Model Application

To address our initial hypothesis that there would be a linear relationship between average entropy and average bout minutes, we created a linear regression model. We plotted the average entropy value and average number of bout minutes for each participant as shown in red below in Figure 5. With 512 data points on the plot, we calculated the line of best fit for the model, as shown in blue. Given how scattered the data is, it was not particularly surprising to find that the R^2 value was 34.64%. As previously described, a higher R^2 value indicates a better fit to the model. Although the data is scattered, it does show a general upward trend indicating that the higher the average

number of bout minutes, the higher the average entropy value and therefore the less habitual participants were.

Figure 5: Linear Regression Model for all Participants

While Figure 5 shows data from the initial 1-year intervention, in Figure 6 we add in the results of the 12-, 18-, and 24-month follow-ups. The plot was created by following the same steps as shown in Figure 5 to find the R^2 values but broken up into 5 milestones: 0 months (baseline), 6 months, 12 months, 18 months, and 24 months. A 7-day sample was used to calculate the R^2 value at each milestone. Since a greater R^2 value indicates a better fit of the data to the model, our model is still showing that the higher the average number of bout minutes, the higher the average entropy value and therefore the less habitual participants were. Here we can see that this is the most accurate for Group A at the 12-month milestone with an R^2 value of 0.4007. We can also see that all groups other than Group A showed an increase in R^2 scores between the 12-month and 18-month milestones, before everyone showed a decrease between the 18-month and 24-month milestones.

Figure 6: R2 Values Including 12-, 18-, and 24-Month Follow-Ups

6 Discussion

The results from our analysis ended up disproving our initial hypothesis that the more consistent a person is about the time of day they exercise, or the more habitual they are, the more they will exercise overall. What we instead uncovered is that people who exercise more are actually less consistent in the time of day that they exercise, which would align with the chaos that most of us experience on a daily basis. In an ideal world, we would all have the ability to take complete control over our lives. If we could plan every minute of our days and follow that exact plan each day without interruption, then achieving our goals would be much easier. Unfortunately, the reality that we do live in does not even come close to allowing for such control. There are so many outside factors that are out of our control, always seeming to pop up at the most inconvenient of times. When we set exact times to consistently exercise, the likelihood of us staying free during that period for any real length of time is extremely low. Then, when we do inevitably become unavailable during that time for a day, we tend to get caught with the mentality of “it’s now or never, so I just won’t exercise today”, and “there’s always tomorrow”. Perhaps if more people had New Year’s resolutions to exercise for 150 minutes every week, fitting it in wherever they happened to have room in their busy schedules, those resolutions might actually last throughout the year, instead of fading away by February.

When comparing the results of the different trial groups, as shown in Figure 6, it’s interesting to see that Group B begins and ends the milestones with the highest R^2 score. Group B had the static physical activity goals, requiring that participants exercise for at least 30 minutes a day for at least 5 days to meet the recommended 150 minutes per week. Participants in Group B also fell into the

immediate financial reinforcement group where points were earned for successfully meeting or exceeding their daily MVPA goal, which could then be exchanged for an e-gift card once they reached 500 points (\$5.00 e-gift card). However, they also faced the unpredictability of the varying reinforcement magnitude and frequency depending on the reinforcement stages. This group may have had the highest R^2 scores due to the mix of structure and predictability that comes with the static goal group paired with the excitement in the unknown from the immediate financial reinforcement group. It could even be argued that participants in the immediate financial reinforcement group may have felt cognitive affects comparable to that of gambling. Studies have shown that gambling impacts both dopamine and serotonin levels in the brain, leading so many to become addicted as a form of self-medication [22, 23]. Although the immediate financial reinforcement group did not have the fear of losing any of their earnings, they still received that comparable excitement of not knowing how much they may earn in any given day, if anything at all. This exhilaration combined with the predictability of having static physical activity goals, may be a combination worth exploring in future studies.

7 Conclusion

With new discoveries and improvements in technology coming out every day with the goal of making our lives easier, our current concerns around a global lack of physical activity are likely to worsen. However, these same changes are also inevitably going to shift our individual goals and motives, making the kind of research discussed in this thesis that much more important. The more we can understand human motivation, especially if the results we find show a contrast in what we may have previously thought to be true, the better we can begin to adapt for success. For example, if the results of this thesis can begin to shift our attitudes away from the idea that we need to exercise consistently at the same time each day to form a healthy habit, and more towards a mindset of flexibility to achieve our goals, we might have a better chance of mitigating the unhealthy trajectory society is currently on. No matter how many people may be reaching for the same goals, they will all have different motives for achieving those goals based on their differing drive, past experiences, environment, and future goals. As our experiences from the current COVID-19 virus pandemic have already begun to show us, flexibility and individuality may be our best foot forward into a successful future.

References

- [1] "Physical Activity," World Health Organization, 2016. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/physical-activity>.
- [2] C. B. Ferster and B. F. Skinner, "Schedules of reinforcement," *Appleton-Century-Crofts.*, Vols. <https://doi.org/10.1037/10627-000>, 1957.
- [3] V. S. Conn, A. R. Hafdahl and D. R. Mehr, "Interventions to Increase Physical Activity Among Healthy Adults: Meta-Analysis of Outcomes," *American Journal of Public Health*, vol. <https://doi.org/10.2105/AJPH.2010.194381>, no. 4, pp. 751-758, 2011.
- [4] S.-H. Chung and R. Herrnstein, "Choice and Delay of Reinforcement," *Journal of the Experimental Analysis of Behavior*, vol. 10, no. 1, pp. 67-74; <https://doi.org/10.1901/jeab.1967.10-67>, 1967.
- [5] J. C. Barte and W. G. Wendel-Vos, "A Systematic Review of Financial Incentives for Physical Activity: The Effects on Physical Activity and Related Outcomes," *Behavioral Medicine*, vol. 43, no. 2, pp. 79-90, DOI: 10.1080/08964289.2015.1074880, 2017.
- [6] S. Skender, J. Ose, J. Chang-Claude, M. Paskow, B. Brühmann, E. M. Siegel, K. Steindorf and C. M. Ulrich, "Accelerometry and physical activity questionnaires - a systematic review," *BMC Public Health*, vol. 16, pp. <https://doi.org/10.1186/s12889-016-3172-0>, 2016.
- [7] M. Adams, J. Hurley, C. Phillips, M. Todd, S. Angadi, V. Berardi, M. Hovell and S. Hooker, "Rationale, design, and baseline characteristics of WalkIT Arizona: A factorial randomized trial testing adaptive goals and financial reinforcement to increase walking across higher and lower walkable neighborhoods," *Contemporary Clinical Trials*, vol. 81, no. 1551-7144, pp. 87-101, (2019).
- [8] Ruescas-Nicolau and Maria-Arantzazu, "Validity of the International Physical Activity Questionnaire Long Form for Assessing Physical Activity and Sedentary Behavior in Subjects with Chronic Stroke," *International journal of environmental research and public health*, Vols. 18,9 4729, 2021.
- [9] B. E. Garfield, J. L. Canavan, C. J. Smith, K. A. Ingram, R. P. Fowler, A. L. Clark, M. I. Polkey and W. D.-C. Man, "Stanford Seven-Day Physical Activity Recall questionnaire in COPD," *European Respiratory Journal*, vol. 40, pp. 356-362, 2012.

- [10] P. S. Freedson, E. Melanson and J. Sirard, "Calibration of the Computer Science and Applications, Inc. accelerometer," *Medicine & Science in Sports & Exercise*, vol. 30, no. 5, pp. 777-781, 1998.
- [11] Diabetes Prevention Program (DPP) Research Group, "The Diabetes Prevention Program (DPP): description of lifestyle intervention," *Diabetes care*, Vols. 25,12, no. doi:10.2337/diacare.25.12.2165, pp. 2165-71, 2002.
- [12] V. Berardi, M. Hovell, J. Hurley, C. Phillips, J. Bellettiere, M. Todd and M. Adams, "Variable Magnitude and Frequency Financial Reinforcement is Effective at Increasing Adults' Free-Living Physical Activity," *Perspectives on Behavior Science*, 43:515-538, (2020).
- [13] M. A. Adams, J. C. Hurley, M. Todd, N. Bhuiyan, C. L. Jarrett, W. J. Tucker, K. E. Hollingshead and S. S. Angadi, "Adaptive goal setting and financial incentives: a 2×2 factorial randomized controlled trial to increase adults' physical activity," *BMC Public Health*, Vols. 17(1), 286, pp. <https://doi.org/10.1186/s12889-017-4197-8>, 2017.
- [14] M. A. Adams, J. F. Sallis, G. J. Norman, M. F. Hovell, E. B. Hekler and E. Perata, "An adaptive physical activity intervention for overweight adults: a randomized controlled trial," *PloS one*, vol. 8(12), no. e82901, p. <https://doi.org/10.1371/journal.pone.0082901>, 2013.
- [15] "R Lubridate Documentation," [Online]. Available: <https://www.rdocumentation.org/packages/lubridate/versions/1.7.4/topics/tz>. [Accessed 2020].
- [16] "Pandas Documentation," [Online]. Available: https://pandas.pydata.org/pandas-docs/stable/user_guide/timeseries.html. [Accessed 2020].
- [17] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, doi: 10.1002/j.1538-7305.1948.tb01338.x, 1948.
- [18] S. I. McClean, "Data Mining and Knowledge Discovery," *Encyclopedia of Physical Science and Technology (Third Edition)*, pp. 229-246, 2003.
- [19] X. Li, J. Yao, X. Liu and H. Guan, "A First Look at Information Entropy-Based Data Pricing," *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, no. doi: 10.1109/ICDCS.2017.45, pp. 2053-2060, 2017.
- [20] "Coefficient of Determination, R-squared," Newcastle University, [Online]. Available: <https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/regression-and-correlation/coefficient-of-determination-r-squared.html>. [Accessed 2021].

- [21] P. Lu, V. Abedi, M. Yongguo, R. Hontecillas, C. Philipson, S. Hoops, A. Carbo and J. Bassaganya-Riera, "Chapter 1 - Supervised Learning with the Artificial Neural Networks Algorithm for Modeling Immune Cell Differentiation," *Emerging Trends in Computer Science and Applied Computing*, no. doi=<https://doi.org/10.1016/B978-0-12-802508-6.00001-6>, pp. 1-18, 2015.

- [22] M. N. Potenza, "The neural bases of cognitive processes in gambling disorder," *Trends in Cognitive Sciences*, vol. 18, no. 8, pp. 429-438; <https://doi.org/10.1016/j.tics.2014.03.007>, 2014.

- [23] D. Campbell-Meiklejohn, J. Wakeley, V. Herbert, J. Cook, P. Scollo, M. Kar Ray, S. Selvaraj, R. E. Passingham, P. Cowen and R. D. Rogers, "Serotonin and dopamine play complementary roles in gambling to recover losses," *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, vol. 36(2), pp. 402-10; doi: 10.1038/npp.2010.170, 2011.