Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Nikita Kovalenko December 6th, 2016

The Effect of White Lies:

How Dishonest Actions Can Subconsciously Influence Self-Concept and Future Decisions

By

Nikita Kovalenko

Kelli Lanier
Adviser

Economics Department

Kelli Lanier

Adviser

Elena Antoniadou

Committee Member

Vera Proskurina

Committee Member

2016

The Effect of White Lies:

How Dishonest Actions Can Subconsciously Influence Self-Concept and Future Decisions

By

Nikita Kovalenko

Kelli Lanier

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics Department

2016

Abstract

The Effect of White Lies:
How Dishonest Actions Can Subconsciously Influence Self-Concept and Future Decisions

By Nikita Kovalenko

People are given opportunities to act dishonestly in their daily lives – how do they generally act? Does being dishonest right now make people less honest later? Multiple studies have shown some patterns in dishonest behavior and that people usually behave a little dishonestly – just enough to where they still feel a strong sense of integrity. This behavior is due to self-concept – our perception of ourselves is a huge influence on how we behave. The perception people have of themselves incorporates a "personal fudge factor" for how much they can cheat without altering how they perceive themselves. By setting up an experiment/exam that allows individuals to act dishonestly without repercussion and actually incentivizes it, we can see this phenomenon at work. The question is, how does what we choose to do one time affect the next? When the same group of students take the exam multiple times without fear of being traced, we get an idea. We find that our "personal fudge factor" needs time to reset itself, for example. Unfortunately, we also find things we don't completely account for which limits the extrapolation of more data. That being said, the framework of this experiment can be built upon to discover just how exactly our morality influences our behavior.

The Effect of White Lies:

How Dishonest Actions Can Subconsciously Influence Self-Concept and Future Decisions

By

Nikita Kovalenko

Kelli Lanier

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Economics Department

2016

Table of Contents

I. Introduction

Economics is the study of human decision-making. Decision-making is a process influenced by various factors – many of which society still has much to learn about, such as morality and self-concept. Evidence suggests that when people occasionally behave dishonestly, their behavior does not alter how they perceive themselves (Ariely et. al., 2006). For example, every year many people partake in return fraud– buying items they intend to return after use. Return fraud makes up around 3.5% to 6.1% of sales and cost companies roughly $15.9 billion in 2015 according to the National Retail Federation (NRF 2015). Other forms of dishonesty may arise in lobbying for political favors – examples being companies like Enron, or individuals like Jack Abramoff. Various other forms of political corruption are well documented and pervasive. The Association of Certified Fraud Examiners estimated employee theft and fraud cost $600 billion a year in the United States alone, a number that well eclipses the total worth of Alphabet, aka Google, (Association of Certified Fraud Examiners 2006)! What is more surprising is that these estimates may be lower than the true costs. If society can better understand the costs of dishonest behavior, effective cost/benefit analysis may be done to understand how active society must be in curbing it. To better estimate the cost of dishonesty, several questions should be answered.

When given multiple opportunities to increase profit by behaving dishonestly, how do people respond? Do they lose some moral self-control and behave more dishonestly, or do they have a limit to how much they can cheat in a set period of time? Regarding the previous examples of dishonesty, the data from this experiment may help us better approximate the frequency of return fraud, political corruption, and employee misconduct. This can in turn

improve the detection process for these monitoring actions by detecting general patterns and possibly predicting future dishonest actions. If it is concluded that cheating increases exponentially over time, estimates of its costs to society may be understated. Hence, society may find benefits of actively attempting to curb dishonest behavior worth the cost.

To determine how small, individual immoral decisions may affect individual decisions in the future, I performed a similar experiment to the one economists in Ariely et. al. 2008 used during their research on a similar topics – a simple matching quiz with treatment groups self-reporting results. Such an experiment, when conducted with some minor adjustments, allowed me to track individual results over time. At the very least, this gave me the opportunity to observe if a rest period existing between opportunities to commit small immoral actions altered behavior.

II. Literature Review

a. How Decisions are influenced by Morality

Over the last several decades, economists have studied morality closely and discovered several interesting behavioral patterns. Nina Mazar, On Amir, and Dan Ariely attempted to quantify cheating and immoral behavior in their 2006 experiment. In this experiment, they issued every subject in two groups a twenty question matching quiz, but gave them only four minutes to answer the questions. One group had their exams collected (the control group), while the other had the students self-report their exam results (treatment group). Subjects were notified beforehand that they would be paid for each question answered correctly and were given reasons to infer they could cheat without ramifications. The results showed the

control group answered an average of about 3 questions correctly each time, while the treatment group claimed to answer roughly 6 (Ariely et. al. 2006). Cheating/immoral behavior was detected, but subjects did not take the most money possible. What was more surprising is that the experimental results were duplicated when the experiment was administered again, only this time with an increase in payout per question from $.50 to $2 (Ariely et. al. 2008). The conclusion from this experiment was that people have a personal "fudge factor" – they can be dishonest up to a certain amount of questions without changing how they perceive themselves – also known as their "self-concept." This conclusion raised multiple questions concerning the specific characteristics of this moral "fudge factor" which others, including myself, aspired to answer.

b. External Influences on Morality

Due to the information gathered from the previous tests, Ariely, Mazar, and Amir conducted several more experiments to examine what influences this personal "fudge factor." They discovered that heightening subjects' awareness of their morality by asking them to either attempt to recite the Ten Commandments or by signing an honor code before the test reduced cheating/immoral behavior (Ariely et. al., 2006). Moreover, they found that separating people from the payout by giving tokens to later exchange for cash increased immoral behavior significantly (Ariely et. al., 2008). In another paper, Francesca Gino, Shahar Ayal and Dan Ariely tested the influence of the "in-group effect" on dishonesty. What they found is that if a member of a subject's "in-group" (i.e. if they attend the same university as the subject in the case of this experiment) acts blatantly immoral, the subject is more likely to behave more immorally. Moreover, the study also consisted of subjects in which the blatantly dishonest

person was not a member of their "in-group," and they acted less immorally (Ariely et. al., 2009). This second experiment also used an important technique to measure cheating/immoral behavior – it provided students with the highest amount of money they could earn within envelopes and instructed students to remove the amount of money they earned at the end of the test themselves. While it is important to know what actions influence morality, it is also important to discover how internal factors influence morality.

c. Internal Influences on Morality

While there are plenty of extrinsic factors at play in morality, there are also several intrinsic factors. Two of these are exhaustion and willpower. Nicole L. Mead et. al. conducted an experiment in which they asked subjects to write a short essay without the letters A and N – an exhaustive task. Subjects with this treatment cheated significantly more when taking the matching quiz immediately after (Mead 2009). Another potentially relevant study is the one by Polivy et. al., which tests the "What the Hell" effect. This test demonstrated that when someone fails at a goal they set for themselves, they generally proceed to continue failing in a more extreme manner (Polivy et. al. 2010). For example, when someone sets a goal to diet and fails, they are more likely to binge eat afterward.

d. Significance to Experiment

The findings of these experiments are significant in how they shaped my question and in how my experiment was designed. The previous studies leave two questions unanswered: 1) how much time is needed for the personal "fudge factor" to completely reset? 2) How do immoral decisions influence future decisions? Furthermore, the previous experiments highlighted factors I needed to control for in my own experiment – such as exhaustion. I believe

these questions are worth exploring, as morality plays a role in everyday decisions, and I am confident that the experiment designed provides some answers.

III. Methodology

a. Subjects and Experimental Setting

For the experiment, I used two separate class sections of "Economics 215: Stocks, Bonds and Financial Markets," as my sample groups. There were several reasons these classes were selected: they provided rosters with adequate sample sizes (n>30), allowed me to control more factors while administering the exam, and theoretically had a strong likelihood of subject return – which is necessary for an experiment conducted over the course of a week.

While the class rosters were larger than 30 students, not all of them attended each class. As a result, my control group and treatment group, after removing students that did not attend each day, had sample sizes of 18 and 25, respectively. However, these students partook in the entirety of the experiment and the samples were large enough to produce statistically substantial results.

The experiment controlled for location and type of student – Emory students studying economics. Using samples from the same course and professor can also control for factors of pressure the subjects may feel before attending the class. The classrooms were the exact same location each day, the difference in administration between the control and treatment group was simply the time of day. The treatment group immediately followed the control group – the control exam was administered at 10am, the treatment at 11am. The time of day was not a significant factor in the other experiments noted earlier.

The subjects were administered 4 tests – 2 on Tuesday, 1 on the Thursday of the same week, and the final on the Friday of that same week. All exams, except for the second, were administered at the beginning of the class. The second exam was administered at the end of the Tuesday class. To avoid potential exhaustion effects, the exam was scheduled the week before the last week of the semester. Students typically have less exams then, as professors tend to give finals during the allotted time at the end of the semester instead. Moreover, the exhaustion effect is accounted for by asking students to fill out a survey at the end of each exam indicating how exhausted they feel on a scale from 1 to 10.

b. Experimental Procedure

Subjects entered the classroom without the professor present. The proctor provided the materials to the students, who then filled out the consent form. Subjects were then given instructions as to how the exam would be completed and how they would be paid. The subjects received 3 minutes to complete as many matrices – which will be described in the following subsection – as they could. In the treatment group, photographs were taken without the subjects' knowledge. Once 3 minutes passed, students were told to stop taking the test and were asked to complete a quick survey. In the control group, students were asked to note a 4 digit identification code that they continued to use on each exam. In the control group, students were then paid in private by the proctor. In the treatment group, subjects were asked to self-report their results by removing money from the envelope beneath their seat. Furthermore, students were instructed to not discuss the experiment with their peers. Once the test was completed, the professor entered the class.

c. Matrix Test

Every subject was asked to take four quizzes, which consisted of 20 three-by-four matrices. Each cell had a number with two decimal places. For each matrix, subjects had to find two numbers that added up to ten. Here is an example of one below:

1.69	1.82	2.91
4.67	4.81	3.05
5.82	5.06	4.28
6.36	5.19	4.57

This task was selected because it is fairly simple to understand and subject performance does not have significant variance. Subjects were given three minutes to complete each question, and were informed prior to the quiz that they would receive $.25 for every matrix they solve.

d. Treatment

The experiment consisted of two groups – a control group and a treatment group. The treatment group was administered the same tests and surveys as the control group. The difference between the groups was the method in which they were compensated. Each subject in the treatment group was provided envelopes with $5, made up of four one dollar bills and four quarters. While taking the test, the classroom was photographed as aforementioned. The subjects were asked to check their work and were instructed to remove only $.25 from the envelope for each question they correctly answered. After taking their earnings, subjects were instructed to leave the envelope below their seats. The amount each subject claimed to have answered correctly was determined by the money left in the envelope when the proctor

collected it. These envelopes were collected in a methodical order to ensure the results corresponded to the correct subject.

e. Survey

Surveys were administered three times throughout the experiment – once after the second, third and fourth tests. Students were asked to complete the surveys after each exam. The first two surveys asked students to rate their exhaustion level and sleep amount. These questions were asked, in part, to deceive the students from guessing the true intentions of the exam. Subjects were asked to write their PIN on their survey in the control group, whereas in the treatment group they were asked to insert the surveys inside their envelopes.

At the end of the experiment, a survey was administered to subjects to collect generic demographic data. No demographic data provided statistically significant results.

f. Hypothesis

As mentioned earlier, the goal of the experiment is to find whether small, individual, immoral decisions affected future behavior. By allowing treatment subjects to self-report their results without any apparent consequence, risk averse individuals were encouraged to cheat. Over time we can compare these results to the control group and determine how self-concept and morality influenced the behavior of the treatment group.

Prior to the experiment I made two hypotheses. The first was that test 1, 3 and 4 medians of the treatment group would be significantly greater than those of the control group, implying the treatment group had cheating members. The idea was that the testing would produce high z-scores. The second hypothesis was that as the experiment continued, cheating within the treatment group would grow exponentially – meaning the median difference

between exams would be greater than 0 for the treatment group and equal 0 for the control group.

IV. Data Analysis

Cheating/immoral behavior was determined by comparing results of the treatment group to the control group using the Mann-Whitney statistical test – a non-parametric exam that can be used for small sample sizes. This test was used because the results showed an abnormal distribution. To track data across time, a Wilcoxon Signed Rank test – also a non-parametric exam – was used, which allowed me to determine individual changes in data over time for both the treatment and the control group. The Wilcoxon Signed Rank test used the difference between the results of each individual between tests, and provided proper weighting to account for the abnormal distribution. The result, ultimately, would show if a significant difference in test taking performance existed.

a. Between Control and Treatment

The results of the experiment were statistically significant in a number of ways and confirmed parts of the hypothesis to be true. Using a Mann-Whitney test, data between tests 1, tests 2, tests 3 and tests 4 in the control and treatment were tested against each other, respectively. The null hypothesis in each statistical test was that the medians of each group were equal. For test 1, $p= 0.000012 < 0.001$, so the null hypothesis that the two groups' medians were same was strongly rejected. Hence in test 1, cheating was statistically detectable. For test 2, $p=0.2835$, the results could not detect that the medians between the control and treatment groups were significantly different. In other words, cheating could not be detected.

For test 3 and 4, p=0.000056 and p=0.000031, both strongly rejecting their respective null hypotheses when alpha=.001. The testing and data can be seen in Appendixes A, B, C, and D. The data suggests that no cheating occurred during the second test, likely due to the subjects' self-concept. An interpretation of these results could be that for one to feel honest despite behaving dishonestly, one needs to wait more than an hour between small dishonest actions.

	Test 1	Test 2	Test 3	Test 4
$Median_c$	3.5	5.5	6	6
$Median_t$	6	6	10	12
m=	18	18	18	18
n=	25	25	25	25
N=	43	43	43	43
Wc=	229.5	373.5	242.5	241.5
Wt=	716.5	572.5	703.5	704.5
z=	4.2219	0.5725	3.8598	4.0087
p=	0.000012**	0.2835	0.000056**	0.000031**
** significant when alpha = .001				

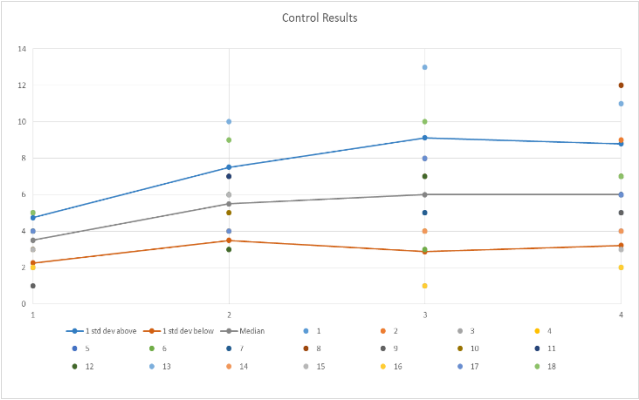
b. Across Treatment and Control

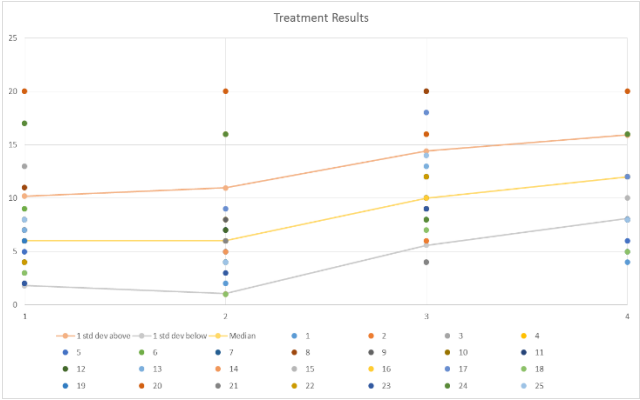
The next part of my analysis involved comparing the performances of the subjects throughout the experiment within their own group. For this, I used a Wilcoxon Signed Rank test, which compares performance across groups. I compared the difference in results between tests 2 and 1, 3 and 2, 4 and 3, and 4 and 1. The null hypothesis for each group was that the medians of the exams were the same.

Starting with the control group, I rejected the null hypothesis when detecting change between exams 2 and 1 and exams 4 and 1. However, I failed to reject the null between tests 2 and 3 and 3 and 4. The control group performed significantly better on the second exam, most

likely due to improved performance in the task. Due to the improvement in performance, it is no surprise the null between tests 1 and 4 was rejected as well.

The treatment group showed different results. For tests 2 vs. 1, the null that the medians were the same failed to be rejected, as self-concept maintenance likely became a factor as to why the results did not differ statistically. For test 3 vs. 2, the null was rejected, meaning self-concept was no longer playing a role in decision-making. An improvement in performance in the control group after the first exam may explain the jump between exams 2 and 3. Tests 3 vs. 4 did not show any difference in medians, and I failed to reject the null. Because of the jump, the results of tests 1 vs. 4 were statistically significant, and I rejected the null. More data regarding this testing can be seen in Appendixes E-L.





re	There is no difference in improvement between tests 1 and 2	There is no difference in improvement between tests 2 and 3	There is no difference in improvement between tests 3 and 4	There is no difference in improvement between tests 1 and 4
Control	Reject the Null, p=0.000017	Fail to Reject the Null, p=0.271028	Fail to Reject the Null, p=0.430853	Reject the Null, p=0.000685
Treatment	Fail to Reject the Null, p=0.221586	Reject the Null, p=0.000901	Fail to Reject the Null, p=0.242095	Reject the Null, p=0.00007

When taking both results of the Mann-Whitney testing and the Wilcoxon Signed Rank testing, a clear picture is painted. Cheating was not detected in the treatment group during the second test because subjects did not have enough time to reset their moral “fudge factors” and

cheat again. However, after 23 hours, subjects were willing to cheat again, as significant improvement was seen in the treatment group, but not the control.

c. Survey Data

After testing between and across the many different demographics, I found only one statistically significant result. Males claimed to answer more matrices than females on the final test of the treatment group. The sample sizes were too small to determine whether socioeconomic status, exhaustion, being raised in urban or suburban areas, perception of others' honesty, perception of the subject's own honesty, and many other factors had an influence on dishonest behavior. To view what was asked in the survey, one can look at Appendix M. Most students in the treatment group claimed that they were very honest individuals (76%), more honest than others in their class (64%), and said they were extremely honest in their result reporting (84%). The perception of self-honesty was even shared by the person in the treatment group who took all the money from the envelope in three of the four exams, and left only the quarters in the other. Statistically, the control group's subjects view of how honestly they viewed themselves and others did not deviate significantly from the treatment group's subjects.

V. Conclusion

The goal of the experiment was to determine whether or not small, individual, immoral decisions influence future moral behavior, and the results were clear – they do. The data demonstrated a strong disparity between the treatment and control groups in three of the four tests. In the second test, cheating could not be detected, strengthening the notion that time

must go by before one typically behaves dishonestly again. Moreover, most individuals claimed they were honest – which is a testament to self-concept maintenance. Very few students cheated at high levels for a long time, signifying that self-concept maintenance remained influential in the decision-making of the subjects.

When discerning whether individuals became more or less dishonest after a day, I found my results to be inconclusive due to the unforeseen jump in improvement in the control group on exam 2. While cheating was still detected in exams 3 and 4, it is difficult to determine whether it was any more or less significant than in the first exam due to the small sample size.

The small sample size may be increased to check the validity of these findings, as well as to determine the impact certain demographic factors may play in decision-making Unfortunately, the sample size was too small to conclude almost anything from the surveys – even the factor of exhaustion which was tested and shown to be significant in another similar study.

VI. General Discussion

Like many behavioral experiments, this one can be improved upon. The small sample size can be increased to check the validity of these findings, as well as determine the impact certain demographic factors may play in this decision. Unfortunately, the sample size was too small to conclude almost anything from the surveys – even the factor of exhaustion which was tested and shown to be significant in another similar study.

The experiment did provide valuable information regarding our self-concept maintenance as previously stated. This must be further explored to determine the average time

a typical individual's morality may need to reset. If the experiment concluded anything, it was that after behaving dishonestly, at least an hour to a day is needed to reset one's "personal fudge factor." Moreover, a larger sample needs to be present in order to find out how different demographics may behave.

Furthermore, to control for the jump in data that prevented me from concluding whether small dishonest actions increase dishonesty in the future, practice rounds of the test should be done beforehand. As the control group's performance became more constant after the performance jump, I believe this measure would help produce meaningful results.

In addition to this, I would like an experiment to be designed to test how people report results after they find out they fell short of what their initial expectations were.

That being said, this experiment can serve as a preliminary study for further honesty experiments.

VII. References

Bibliography

Ariely, Dan, Shahar Ayal, and Francesea Gino. "Contagion and Differentiation in Unethical Behavior: The Effect of One Bad Apple on the Barrel." Psychological Science 20.3 (2009): 393-98. Web.

Ariely, Dan and Nina Manzar. "Dishonesty in Everyday Life and Its Policy Implications." Journal of Public Policy & Marketing 25.1, Helping Consumers Help Themselves: Improving the Quality of Judgments and Choices (2006): 117-26. Web.

Ariely, Dan, Nina Manzar, and On Amir. "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." Journal of Marketing Research 45.6 (2008): 633-44. Web.

Mead, Nicole L., Roy F. Baumeister, Francesca Gino, Maurice E. Schweitzer, and Dan Ariely. "Too Tired to Tell the Truth: Self-control Resource Depletion and Dishonesty." Journal of Experimental Social Psychology 45.3 (2009): 594-97. Web.

Polivy, Janet, C. Peter Herman, and Rajbir Deo. "Getting a Bigger Slice of the Pie. Effects on Eating and Emotion in Restrained and Unrestrained Eaters." Appetite 55.3 (2010): 426-30. Web.

Powers, Rod. "Consumer Returns in the Retail Industry." Online Posting. National Retail Federation. NRF, Dec. 2015. Web. 27 Nov. 2016.

Report to the Nation on Occupational Fraud and Abuse. Austin: ACFE, 2006. Association of Certified Fraud Examiners, 2006. Web. 27 Nov. 2016.

VIII. Appendix

A. Mann-Whitney Test 1

rank	score	group
1	1	C
4	2	T
4	2	C
4	2	C
4	2	C
4	2	C
9.5	3	T
9.5	3	T
9.5	3	C
9.5	3	C
9.5	3	C
9.5	3	C
16.5	4	T
16.5	4	T
16.5	4	T
16.5	4	C
16.5	4	C
16.5	4	C
16.5	4	C
16.5	4	C
23	5	T
23	5	C
23	5	C
23	5	C
23	5	C
28.5	6	T
28.5	6	T
28.5	6	T
28.5	6	T
28.5	6	T
28.5	6	T
33	7	T
33	7	T
33	7	T
36.5	8	T
36.5	8	T
36.5	8	T
36.5	8	T
39	9	T
40	11	T
41	13	T
42	17	T
43	20	T

n	m	N
18	25	43

null	control and treatment have the same distribution
alternative	Treatment is stochastically larger than Control
null P[T > C] = 1/2	
accept alternative if P[T > C] > 1/2	
= median of T > median of C	

rank sum C:	Wc =	229.5			
rank sum T:	Wt =	716.5			

z=(Wt +- .5 - m*(N+1)/2)/(sqrt([m*n/(N*(N-1))][(N^3-N)/12 - sigma(tn^3-tn)/12])

z=	4.221855
p=	0.000012

B. Mann-Whitney Test 2

rank	score	group
1.5	1	T
1.5	1	T
3	2	T
5	3	C
5	3	C
5	3	T
12	4	C
12	4	C
12	4	C
12	4	C
12	4	C
12	4	C
12	4	T
12	4	T
12	4	T
12	4	T
12	4	T
19	5	C
19	5	T
19	5	T
24	6	C
24	6	C
24	6	C
24	6	T
24	6	T
24	6	T
24	6	T
31	7	C
31	7	C
31	7	C
31	7	C
31	7	T
31	7	T
31	7	T
35.5	8	T
35.5	8	T
37.5	9	C
37.5	9	T
39	10	C
41	16	T
41	16	T
41	16	T
43	20	T

n	m	N
18	25	43

null	control and treatment have the same distribution
alternative	Treatment is stochastically larger than Control
null P[T > C] = 1/2	
accept alternative if P[T > C] > 1/2	
= median of T > median of C	

rank sum C:	Wc =	373.5			
rank sum T:	Wt =	572.5			

z=(Wt +- .5 - m*(N+1)/2)/(sqrt([m*n/(N*(N-1))][(N^3-N)/12 - sigma(tn^3-tn)/12])

z=	0.57254	0.598565
p=	0.283478	0.274732

C. Mann-Whitney Test 3

rank	score	group
1.5	1	C
1.5	1	C
3.5	3	C
3.5	3	C
7	4	C
7	4	C
7	4	C
7	4	T
7	4	T
10.5	5	C
10.5	5	C
12	6	T
14.5	7	C
14.5	7	C
14.5	7	C
14.5	7	T
20	8	C
20	8	C
20	8	C
20	8	C
20	8	T
20	8	T
20	8	T
26	9	T
26	9	T
26	9	T
26	9	T
26	9	T
30.5	10	C
30.5	10	T
30.5	10	T
30.5	10	T
34	12	T
34	12	T
34	12	T
36.5	13	C
36.5	13	T
38	14	T
39.5	16	T
39.5	16	T
41	18	T
42.5	20	T
42.5	20	T

n	m	N
18	25	43

null	control and treatment have the same distribution
alternative	Treatment is stochastically larger than Control
null P[T > C] = 1/2	
accept alternative if P[T > C] > 1/2	
= median of T > median of C	

rank sum C:	Wc =	242.5			
rank sum T:	Wt =	703.5			

z=(Wt +- .5 - m*(N+1)/2)/(sqrt([m*n/(N*(N-1))][(N^3-N)/12 - sigma(tn^3-tn)/12])

z=	4.221855	4.247288
p=	0.000012	0.000011

D. Mann-Whitney Test 4

rank	score	group
1.5	2	C
1.5	2	C
3.5	3	C
3.5	3	C
6	4	C
6	4	C
6	4	T
9.5	5	C
9.5	5	C
9.5	5	T
9.5	5	T
14	6	C
14	6	C
14	6	C
14	6	C
14	6	T
18	7	C
18	7	C
18	7	C
23	8	T
23	8	T
23	8	T
23	8	T
23	8	T
23	8	T
23	8	T
23	8	T
27	9	C
28	10	T
29	11	C
34.5	12	C
34.5	12	T
34.5	12	T
34.5	12	T
34.5	12	T
34.5	12	T
34.5	12	T
34.5	12	T
34.5	12	T
34.5	12	T
41	16	T
41	16	T
41	16	T
43	20	T

n	m	N
18	25	43

null	control and treatment have the same distribution
alternative	Treatment is stochastically larger than Control
null P[T > C] = 1/2	
accept alternative if P[T > C] > 1/2	
= median of T > median of C	

rank sum C:	Wc =	241.5			
rank sum T:	Wt =	704.5			

z=(Wt +- .5 - m*(N+1)/2)/(sqrt([m*n/N*(N-1)][(N^3-N)/12 - sigma(tn^3-tn)/12])

z=	4.008686	4.034717
p=	0.000031	0.000027

E. Wilcoxon Signed Rank Control Test 1

Subject	Test 1	Test 2	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24					
17	4	4	0	1	0									
2	5	4	-1	3	1	-			87.5					
12	4	3	-1	3	1	-			21.12463					
6	2	3	1	3	1	+		z=	4.142084					
3	5	7	2	7.5	2	+		p=	0.000017		a=.01			
4	4	6	2	7.5	2	+								
5	2	4	2	7.5	2	+		Ho:	There is no difference between performance in test 1 and 2					
10	3	5	2	7.5	2	+		Ha:	There was improvement in performance from 1 to 2					
14	2	4	2	7.5	2	+								
16	2	4	2	7.5	2	+		Reject the null						
1	4	7	3	13	3	+								
7	3	6	3	13	3	+								
9	1	4	3	13	3	+								
11	4	7	3	13	3	+								
15	3	6	3	13	3	+								
8	3	7	4	16.5	4	+								
18	5	9	4	16.5	4	+								
13	5	10	5	18	5	+								
n=	18		T+=	164		T-=	6							
n>15			N=	17			0 is omitted							

F. Wilcoxon Signed Rank Control Test 2

Subject	Test 2	Test 3	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24					
1	7	7	0	2	0	-								
6	3	3	0	2	0	+			14					
14	4	4	0	2	0	-			22.96193					
5	6	5	-1	6.5	1	-		z=	0.609705					
7	5	4	-1	6.5	1	-		p=	0.271028		a=.01			
8	4	5	1	6.5	1	+								
10	7	8	1	6.5	1	+		Ho:	There is no difference between performance in test 1 and 2					
11	7	8	1	6.5	1	+		Ha:	There was improvement in performance from 1 to 2					
18	9	10	1	6.5	1	+								
4	6	4	-2	10.5	2	-		Fail to reject the null						
15	6	8	2	10.5	2	+								
2	4	1	-3	13.5	3	-								
9	4	1	-3	13.5	3	-								
13	4	7	3	13.5	3	+								
16	10	13	3	13.5	3	+								
3	7	3	-4	17	4	-								
12	3	7	4	17	4	+								
17	4	8	4	17	4	+								
n=	18		T+=	99.5		T-=	71.5							
n>15			N=	18			0 is randomly assigned +/-							

G. Wilcoxon Signed Rank Control Test 3

Subject	Test 3	Test 4	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24				
3	3	3	0	2	0	-							
12	7	7	0	2	0	+			-4				
14	4	4	0	2	0	-			22.96193				
5	5	4	-1	5.5	1	-		z=	-0.1742				
6	3	2	-1	5.5	1	-		p=	0.430853		a=.01		
7	5	6	1	5.5	1	+							
16	1	2	1	5.5	1	+		Ho:	There is no difference between performance in test 1 and 2				
1	7	5	-2	10.5	2	-		Ha:	There was improvement in performance from 1 to 2				
2	7	9	2	10.5	2	+							
4	4	6	2	10.5	2	+		Fail to reject the null					
11	8	6	-2	10.5	2	-							
13	13	11	-2	10.5	2	-							
17	8	6	-2	10.5	2	-							
10	4	7	3	14.5	3	+							
18	10	7	-3	14.5	3	-							
8	8	12	4	16.5	4	+							
9	1	5	4	16.5	4	+							
15	8	3	-5	18	5	-							
n=	18		T+=	81.5		T-=	89.5						
n>15			N=	18			0 is randomly assigned +/-						

H. Wilcoxon Signed Rank Control Test 4

	Test 1	Test 4	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24				
6	2	2	0	2	0	-							
15	3	3	0	2	0	+			73.5				
16	2	2	0	2	0	-			22.96193				
1	4	5	1	4	1	+		z=	3.200951				
3	5	3	-2	8	2	-		p=	0.000685		a=.01		
4	4	6	2	8	2	+							
5	2	4	2	8	2	+		Ho:	There is no difference between performance in test 1 and 2				
11	4	6	2	8	2	+		Ha:	There was improvement in performance from 1 to 2				
14	2	4	2	8	2	+							
17	4	6	2	8	2	+		Reject the null					
18	5	7	2	8	2	+							
7	3	6	3	12.5	3	+							
12	4	7	3	12.5	3	+							
2	5	9	4	15	4	+							
9	1	5	4	15	4	+							
10	3	7	4	15	4	+							
13	5	11	6	17	6	+							
8	3	12	9	18	9	+							
n=	18		T+=	159		T-=	12						
n>15			N=	18			0 is assigned +/- randomly						

I. Wilcoxon Signed Rank Treatment Test 1

Subject	Test 1	Test 2	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24					
15	6	6	0	2	0	-								
16	6	6	0	2	0	+			-28.5	28.5				
20	20	20	0	2	0	-			37.16517					
1	3	2	-1	8.5	1	-		z=	-0.76685	0.766847				
7	4	5	1	8.5	1	+		p=	0.221586		a=.01			
9	7	8	1	8.5	1	+								
11	6	7	1	8.5	1	+		Ho:	There is no difference between performance in test 1 and 2					
12	6	7	1	8.5	1	+		Ha:	There was improvement in performance from 1 to 2					
13	7	6	-1	8.5	1	-								
14	6	5	-1	8.5	1	-		Fail to reject the null						
17	8	9	1	8.5	1	+								
23	2	3	1	8.5	1	+								
24	17	16	-1	8.5	1	-								
5	5	7	2	15.5	2	+								
18	3	1	-2	15.5	2	-								
19	6	4	-2	15.5	2	-								
21	8	6	-2	15.5	2	-								
4	4	8	4	19	4	+								
10	8	4	-4	19	4	-								
25	8	4	-4	19	4	-								
6	9	4	-5	21.5	5	-								
8	11	16	5	21.5	5	+								
2	7	1	-6	23	6	-								
3	13	4	-9	24	9	-								
22	4	16	12	25	12	+								
n=	25		T+=	134		T-=	191							
n>15			N=	25			0 is randomly assigned							

J. Wilcoxon Signed Rank Treatment Test 2

Subject	Test 2	Test 3	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24					
5	7	9	2	3.5	2	+								
9	8	10	2	3.5	2	+			116					
11	7	9	2	3.5	2	+			37.16517					
12	7	9	2	3.5	2	+		z=	3.121202					
15	6	4	-2	3.5	2	-		p=		a=.01				
21	6	4	-2	3.5	2	-								
6	4	8	4	9.5	4	+		Ho:	There is no difference between performance in test 1 and 2					
8	16	20	4	9.5	4	+		Ha:	There was improvement in performance from 1 to 2					
14	5	9	4	9.5	4	+								
16	6	10	4	9.5	4	+		Reject the Null						
20	20	16	-4	9.5	4	-								
22	16	12	-4	9.5	4	-								
2	1	6	5	13.5	5	+								
7	5	10	5	13.5	5	+								
1	2	8	6	16	6	+								
18	1	7	6	16	6	+								
23	3	9	6	16	6	+								
13	6	13	7	18	7	+								
4	8	16	8	20.5	8	+								
10	4	12	8	20.5	8	+								
19	4	12	8	20.5	8	+								
24	16	8	-8	20.5	8	-								
17	9	18	9	23	9	+								
25	4	14	10	24	10	+								
3	4	20	16	25	16	+								
n=	25		T+=	278.5		T-=	46.5							
n>15			N=	25										

K. Wilcoxon Signed Rank Treatment Test 3

Subject	Test 3	Test 4	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24					
6	8	8	0	1.5	0	-								
10	12	12	0	1.5	0	+			-26					
11	9	8	-1	4.5	1	-			37.16517					
12	9	8	-1	4.5	1	-		z=	-0.69958					
13	13	12	-1	4.5	1	-		p=	0.242095		a=.01			
23	9	8	-1	4.5	1	-								
2	6	8	2	8.5	2	+		Ho:	There is no difference between performance in test 1 and 2					
9	10	12	2	8.5	2	+		Ha:	There was improvement in performance from 1 to 2					
16	10	12	2	8.5	2	+								
18	7	5	-2	8.5	2	-		Fail to reject the null						
5	9	6	-3	11.5	3	-								
14	9	12	3	11.5	3	+								
1	8	4	-4	15	4	-								
4	16	12	-4	15	4	-								
19	12	8	-4	15	4	-								
20	16	20	4	15	4	+								
22	12	16	4	15	4	+								
7	10	5	-5	18	5	-								
15	4	10	6	20	6	+								
17	18	12	-6	20	6	-								
25	14	8	-6	20	6	-								
3	20	12	-8	23	8	-								
8	20	12	-8	23	8	-								
24	8	16	8	23	8	+								
21	4	16	12	25	12	+								
n=	25		T+=	136.5		T-=	188.5							
n>15			N=	25			0 is assigned +/- randomly							

L. Wilcoxon Signed Rank Treatment Test 4

Subject	Test 1	Test 4	Difference	Rank		Sign		z=	(T+ - n(n+1)/4)/(sqrt((n*(n+1)*(2n+1))/24					
20	20	20	0	1.5	0	-								
25	8	8	0	1.5	0	+			141.5					
1	3	4	1	6.5	1	+			37.16517					
2	7	8	1	6.5	1	+		z=	3.807328					
3	13	12	-1	6.5	1	-		p=	0.00007		a=.01			
5	5	6	1	6.5	1	+								
6	9	8	-1	6.5	1	-		Ho:	There is no difference between performance in test 1 and 2					
7	4	5	1	6.5	1	+		Ha:	There was improvement in performance from 1 to 2					
8	11	12	1	6.5	1	+								
24	17	16	-1	6.5	1	-		Reject the null						
11	6	8	2	12.5	2	+								
12	6	8	2	12.5	2	+								
18	3	5	2	12.5	2	+								
19	6	8	2	12.5	2	+								
10	8	12	4	16	4	+								
15	6	10	4	16	4	+								
17	8	12	4	16	4	+								
9	7	12	5	18.5	5	+								
13	7	12	5	18.5	5	+								
14	6	12	6	21	6	+								
16	6	12	6	21	6	+								
23	2	8	6	21	6	+								
4	4	12	8	23.5	8	+								
21	8	16	8	23.5	8	+								
22	4	16	12	25	12	+								
n=	25		T+=	304		T-=	21							
n>15			N=	25			0 is assigned +/- randomly							

M. Demographic Survey

Survey 3

Please answer all of the questions accurately. Circle only one answer unless instructed otherwise. Your answers are confidential.

1) How exhausted would you rate yourself on a scale from 1-10 (1 = not very, 10 = incredibly)

1 2 3 4 5 6 7 8 9 10

2) How many hours of sleep did you get last night?

0-2 hours 2-4 hours 4-6 hours 6-8 hours 8-10 hours 10 +hours

3) What is your gender?

M F

4) What is your family's socioeconomic status?

Lower Lower-Middle Middle Upper-Middle Upper

5) How often do you exercise? (1 = never, 10= very often)

1 2 3 4 5 6 7 8 9 10

6) What is you graduation year?

2016 2017 2018 2019 2020

7) How would you describe the area you grew up in?

Rural Area Suburban Area Urban Area

8) How creative are you? (1-10)

1 2 3 4 5 6 7 8 9 10

9) Do you agree with the following statement: I am an honest person (1= never, 10= very easily)

1 2 3 4 5 6 7 8 9 10

10) Family income level?

0-25,000 25,000-50,000 50,000-75,000 75,000-100,000 100,000&up

11) Do you attend a religious service, and if so, how frequent?

Never 1-4 times a year Once a month Every other week Every week More

12) Do you agree with the following statement: I get stressed easily (1=not at all, 10-very)

1 2 3 4 5 6 7 8 9 10

13) How honestly did you report your results? (1 = not at all, 10 = very)

1 2 3 4 5 6 7 8 9 10

14) Did you have an exam this week? (Circle all that apply)

No Yes, Monday Yes, Tuesday Yes, Wednesday Yes, Thursday Yes, Friday

15) To what level do your religious beliefs affect daily life? (1=never, 10=always)

1 2 3 4 5 6 7 8 9 10

16) Have you failed any goals you've set?

No Yes, within 2 weeks Yes, within 1 week

17) Did you discuss this experiment with other students?

Yes No

18) How honestly do you think other people did? (1= not at all, 10 = very)

1 2 3 4 5 6 7 8 9 10

19) Have you ever been accused of academic dishonesty?

Yes No

20) What do you think the purpose of the experiment was, and when did you conclude this?