

W241 Final Project: Subliminal Investigation

Jacquie Nesbitt, Amit Karandikar, Toby Petty, Judge Hiciano, Sarah Yifei Wang

12/10/2021

Abstract

Everyone at some point in their lives has had to take a test or assessment. Whether it was for school, interviewing for a job, or even obtaining that new job certification. Most individuals want to perform well on the test, especially when the outcome can provide access to new opportunities or even lead to a material benefit such as a raise or a new job. What if you could do slightly better in these scenarios, than you could have before, without any extra studying, simply by reading a short message? In our experiment, we investigate this question.

We recruited participants via scheduled LinkedIn posts tagged with data science hashtags. The participants were randomly assigned to either the treatment group, where they were primed by a message asking them to imagine they are an expert in the field before taking a data science related quiz, or were assigned to the control group where they were presented a message of the same length without subliminal priming before answering the same questions. The results did not identify any differences with subliminal priming.

Background

Priming is the way our brains process certain stimuli and impulses subconsciously. These stimuli can be words, like affirmations, pictures or scents, or anything that can be perceived with senses. Subliminal priming occurs when an individual is aware they have received certain stimuli but not aware of the effects those stimuli have on them.

Experiments with subliminal priming have been carried out since the 1950s, some of the more recent and notable experiments include the “Apple makes you think different” experiment published in the 2008 Journal of Consumer research, where 341 university students were asked to perform a visual acuity task during which Apple and IBM logos were flashed very quickly. Subjects subliminally exposed to the Apple logo were able to come up with more creative solutions than those who were primed with the IBM logo.

Another particularly relevant example was a Stanford experiment which examined subliminal priming effects on students. Seventy participants were randomly assigned to the treatment group with a computer that flashed words on the screen related to intelligence, or the control group with words unrelated to intelligence, before completing a practice exam. The practice exams were completed before an actual course midterm. It turned out that the intelligence primes increased performance on the midterm compared to neutral primes. What if we are able to use subliminal priming to improve the quality of one’s work?

Hypothesis

Given the studies mentioned in the background section, the question we attempt to answer is: “Can subliminal priming messages improve the quality of work?” We hypothesize that individuals who receive the

treatment subliminal priming message will score 1 point higher on average on the quiz than individuals in the control group who receive the control message. We are focusing on a 1 point increase because in instances like a job interview it is more important to do relatively better than the person before you than it is to be perfect.

The potential outcomes for each individual in the experiment can be stated as:

- Y_0 = participant's quiz score when they received the control message.
- Y_1 = participant's quiz score when they received the treatment message.

Therefore we state our hypothesis as being that the average treatment effect $E[Y_1 - Y_0]$ is positive 1, meaning that those in the treatment group score 1 point higher on average than those in the control group.

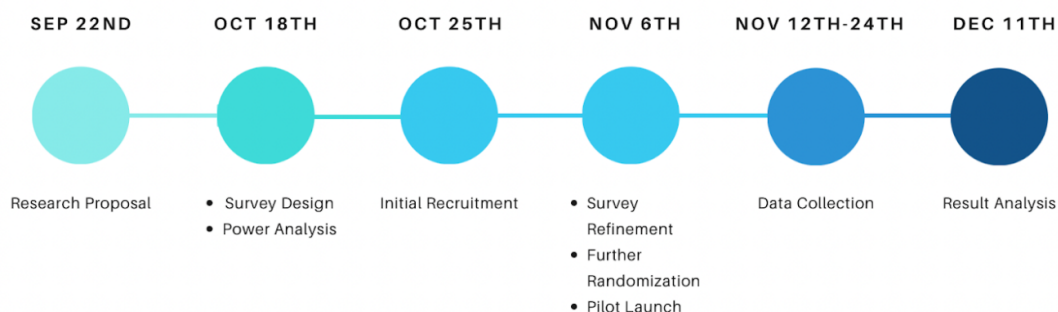
Experiment Design

Experiment Overview

To test our hypothesis we designed a between-subjects experiment. Leveraging Qualtrics, we created a survey that could be deployed via a URL posted on LinkedIn to data scientists, or people with an interest in data science. After answering a few screener questions, we enacted our treatment by presenting the participant a message before receiving an identical data science quiz to solve. The full details of the treatment and control messages are given in the appendix. The problems in the quiz covered a range of subjects, including machine learning concepts, statistics, visualization, computing, etc. We chose to operationalize the quality of work in a quiz format because it can be easily measured numerically. Attention check questions were given after the treatment and control messages to determine if the participants had read the message before starting the quiz questions. Through our power calculation, we determined we needed 126 participants to identify the hypothesized effect size.

Project Timeline

The experiment timeline over 14 weeks is summarized below.



Tools

In this study we used Qualtrics as the main tool for the survey. Qualtrics is web-based software that enables subjects to participate in the survey using a variety of distribution channels, such as phone, laptop, etc.

We leveraged Qualtrics for randomization where Qualtrics randomly assigned subjects into either treatment group or control group with 50%/50% probability in each. Given that each question in the survey had one correct answer, Qualtrics also allowed us to automatically calculate each subject's final score on the quiz.

Power Analysis

$H_0: \mu_t = \mu_c$ - There is no significant difference in the average survey quizlet score between the control and the treatment groups.

$H_a: \mu_t \neq \mu_c$

Using a two-tailed test:

```
pwr.t.test(sig.level = 0.05,
            alternative = "two.sided",
            power = 0.80,
            d = .5)

##
##      Two-sample t test power calculation
##
##              n = 63.76561
##              d = 0.5
##      sig.level = 0.05
##      power = 0.8
##      alternative = two.sided
##
## NOTE: n is number in *each* group
```

Our initial power calculation was based off estimates which we identify as a potential factor in the outcome of our experiment as we used a standard effect size (Cohen's d) of .5 which means we're expecting the average score of the treatment group to be .5 standard deviations away from the control group. In our actual experiment we had 218 participants in our smallest group, therefore we would be able to detect any effects .275 standard deviations and above.

Enrollment and Recruiting

The original design included a plan to send out the survey to the LinkedIn groups "Data Mining, Statistics, Big Data, Data Visualization, AI, Machine Learning, and Data Science" and "Machine Learning and Data Science", both of which have roughly 200,000+ members around the world, offering participants the opportunity to win 1 of 8 \$25 US Amazon gift cards to induce participation. On conducting a pilot survey using this channel however we found a low response rate, and additionally encountered resistance from some of the administrators responsible for those pages.

Subsequently, for the full experiment, we opted for a different strategy on LinkedIn, which was to have 4 members of the research team post a personal message on LinkedIn requesting participation in the survey. To boost the response rate we employed the use of hashtags such as #machinelearning, which has almost 2 million followers on the LinkedIn platform. We also explicitly mentioned that the survey was part of a statistics project to try to induce more responses from those interested in statistics and data science. A sample of the LinkedIn post is shown below:

As part of my statistics course this semester in UC Berkeley's Data Science Master's program my team has to collect our own survey data. If you would like to participate it would be greatly appreciated!

If you complete the survey you'll be entered into a lottery for a chance to win 1 of 8 \$25 Amazon gift cards.

Click the link below to start the survey:

#datascience #machinelearning #ai

Four members of the research team posted the above message with slightly different wording to avoid being flagged as spam, over a period of 12 days from November 12th to November 24th, 2021, at intervals of 3 days to 1 week and at different times of the day, to ensure the post was shown on the hashtags on a variety of weekdays and times to maximize exposure.

Randomization

For randomization we used the Qualtrics randomization feature to randomly assign survey participants to either treatment or control with equal probability. Since our survey questions were focused on data science concepts, we chose to employ a block randomization scheme in which participants were assigned to different blocks based on their data science experience, and then randomized to be in either treatment and control. To assess a participant's skill level in data science, we opted to use a wide-ranging question based on the amount of time they had spent engaged in any data science related activities, as follows:

How many years have you been actively engaged in activities related to learning or practicing data science concepts? Such activities can include: studying either in school/university or through online courses, working in a data science related role, or participating in data science competitions.

The answers available to choose from were:

- Less than 2 years
- More than 2 but less than 5 years
- More than 5 but less than 10 years
- More than 10 years

We reasoned that since there are many different ways in which a person can engage and develop skills in data science, we needed a measure which encompassed all those different activities to create the blocks, and made the simplifying assumption that the longer a person has spent engaged in any data science practice, the more likely they are to have been exposed to the range of concepts which were covered in the survey, and therefore the more likely they were to achieve a better survey score.

Once respondents had answered the question above, those who selected one of the first 2 options were put into a single block for randomization (referred to as “novices”), and those who selected the third or fourth option were put into a second block (referred to as “experts”). Participants in each block were then randomized with a 50/50 probability of being assigned to treatment or control. We used pure random assignment in each block, so that we were not guaranteed to have exactly 50% of participants in treatment/control, but that with enough participation we would approach an even split.

Covariate Balance Check

In order to assess whether our randomization scheme was successful in producing groups which were otherwise random apart from their level of experience in data science, we performed a covariate balance check to assess whether there were any significant differences between treatment and control in any of the other covariates we collected. We regress assigned treatment using just the blocking variable (short model), and then using the other covariates we collected in the survey (long model), and then calculate the F-Statistic. The results are shown below:

```

short_model <- d_without_attrition[,lm(subliminal ~ experience_ind )]
long_model <- d_without_attrition[,lm(subliminal ~ experience_ind +
                                     younger_age + as.factor(education) )]
ftest = anova(short_model, long_model, test='F')
ftest

```

```

## Analysis of Variance Table
##
## Model 1: subliminal ~ experience_ind
## Model 2: subliminal ~ experience_ind + younger_age + as.factor(education)
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1     459 114.87
## 2     454 114.34   5   0.52787 0.4192 0.8354

```

With a P-Value of 0.835422, the long model fails to reject the null hypothesis, we can say that our covariates do not help predict treatment assignment. Therefore, we conclude it is likely there were no significant problems with the randomization scheme.

In our chosen survey design we opted to include the covariate questions at the end of the survey, after the data science related questions which were scored. The reasons for this were based on concerns that if people were asked demographic questions such as age, etc. at the start of the survey they might decide to quit before taking the test, and we wanted to maximize the sample size to obtain the best possible estimate of the treatment effect (especially after the pilot survey in which we found low levels of participation).

Observations and Outcome of Interest

Each participant who chose to complete our survey was asked to answer the same set of questions after they were assigned to either the treatment or control group. The full survey consisted of the following question sections:

1. Introduction, country of residence, and screener question to check the participant had at least some knowledge of what data science is.
2. Blocking question about data science experience (described in the *Randomization* section above).
3. Treatment/Control message (see Appendix).
4. Data science quiz - 10 questions covering the topics of statistics, R code, mathematics, data engineering, machine learning, data visualization libraries, and Big O complexity of Python code. All participants in both the treatment and control group received identical questions.
5. Post-quiz questions asking for demographic information, highest level of completed education, and whether or not they consulted external reference sources in order to answer the quiz.
6. Feedback box and option to enter email address for chance to win the gift card.

The outcome measure determined from the above survey structure was the total number of correct answers supplied to the 10 data science quiz questions. Each question in the quiz had only a single correct response, only 1 selection was allowed of 4 possible answers in each question (i.e. not multiple choice), and not selecting an answer was also an option. Therefore if all participants were to randomly guess responses instead of thinking about the questions, the expected average score would be $\frac{1}{5} \cdot 10 = 2$. Given this expected “random answer” average score of 2 we first tested to see if the scores achieved by participants were significantly better or worse than random:

All participants

T-tests to check if the scores achieved are better than would be expected if participants just randomly guessed.

```
t.test(d$score, mu = 2)
```

```
##
## One Sample t-test
##
## data: d$score
## t = 17.335, df = 463, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 2
## 95 percent confidence interval:
## 3.026135 3.288520
## sample estimates:
## mean of x
## 3.157328
```

```
t.test(d[d$subliminal == 0]$score, mu = 2)
```

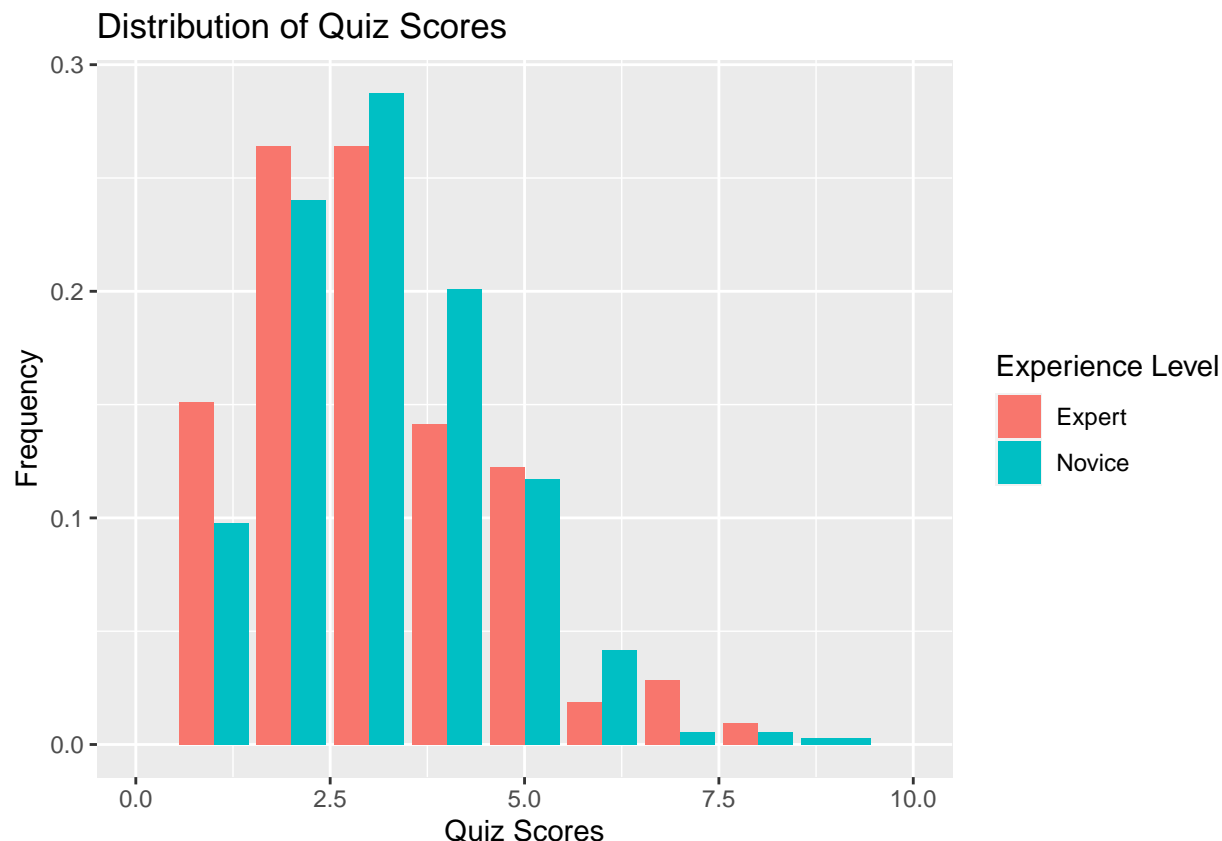
```
##
## One Sample t-test
##
## data: d[d$subliminal == 0]$score
## t = 13.118, df = 217, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 2
## 95 percent confidence interval:
## 3.064137 3.440450
## sample estimates:
## mean of x
## 3.252294
```

```
t.test(d[d$subliminal == 1]$score, mu = 2)
```

```
##
## One Sample t-test
##
## data: d[d$subliminal == 1]$score
## t = 11.524, df = 245, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 2
## 95 percent confidence interval:
## 2.889742 3.256599
## sample estimates:
## mean of x
## 3.073171
```

The t-tests above for all participants, the control group, and the treatment group, all show that performance on the quiz was highly statistically significantly better than would be expected from random guessing. In the graph below, we see that the distribution of scores for *novices* peaks around a score of 3, while the distribution for *experts* is highest between a score of 2 and 3. From this we conclude that on average,

participants did at least try to perform well on the quiz, and did have some knowledge of the data science concepts being tested.



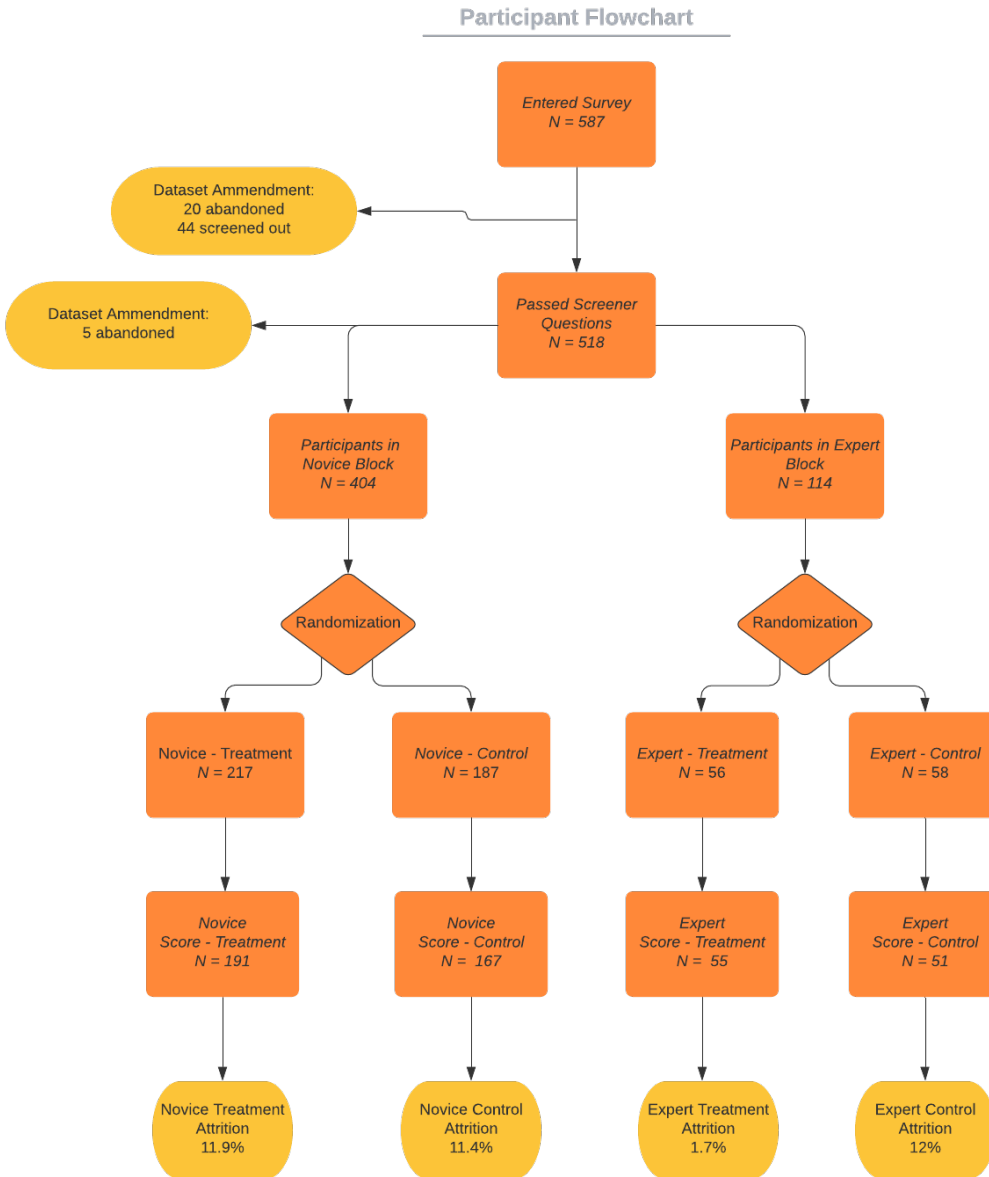
Data Quality, Attrition, Noncompliance, Spillover

Data Quality

One important issue we found with data quality is that Qualtrics indicated that in our survey we had 28 observations that had a high likelihood of being a bot. Also, even though we enabled Qualtrics’ “Prevent Ballot Stuffing” feature which is supposed to prevent people from taking the survey multiple times, it appears that some individuals bypassed this measure, presumably to increase their chances of winning one of the gift card lottery prizes. However, since these observations were subject to the randomization scheme we chose to keep them in for the analysis, since these entries would be equally likely to be assigned to treatment or control. However we recognize that this has implications for our results, since it reduces the statistical power of our experiment to identify a treatment effect by adding more noise to the observations.

Consort Flow Diagram

As seen in the flowchart below, 587 people entered the survey, however, only 523 passed the screener questions. Additionally, 5 abandoned before being randomized on our instructions page. From there, 404 participants were randomized in the *novice* block and 114 randomized in the *expert* block.



Attrition

We define attrition as those who were randomized in their block, but did not complete the quiz. The *novice* block had 217 in treatment and 187 in control, with 11.9% attrition for treatment and 11.4% attrition for control. The *expert* block had 56 in treatment and 58 in control, with 1.7% attrition in treatment and 12% attrition in control.

We were concerned that there might be differential attrition. Therefore, we ran a linear regression seeing if treatment could predict missing scores. According to the regression the treatment is near 0 and not statistically significant, therefore we have determined there is no differential attrition.

Table 1:

	<i>Dependent variable:</i>
	Missing Quiz Score
subliminal	-0.013 (0.027)
experience indicator	-0.044 (0.028)
Constant	0.121*** (0.021)
Observations	518
R ²	0.004
Adjusted R ²	0.0001
Residual Std. Error	0.306 (df = 515)
F Statistic	1.018 (df = 2; 515)
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Noncompliance

We define compliance as an individual who has read the passage carefully. However, because we have incomplete control over whether a participant carefully read the passage, we report our results as intent to treat effects.

To gain a window into possible noncompliance the team implemented the use of attention check questions after the treatment and control messages. Unfortunately, these questions were too vague to act as attention check questions. This could have been remedied by including a phrase before each question stating “Using the message above please answer the following questions”. This would have made it clear that we were asking questions related to the treatment/control message and not their personal opinion.

Additionally, after reviewing the answer choices for the attention check questions we noticed there were several instances where another answer could easily be a logically correct answer instead of the one the team deemed as correct. For example, we ask: “What makes you competitive in the field?”

The answer choices were:

- A. You work out a lot and started training at young age.
- B. You have a strong combination of scientific background, and computational and analytical skills.
- C. You are rich.

The answer the team deemed as correct would be answer choice B. However, answer choice A could also be considered a plausible answer. Answer choice A could also be understood as “someone who works a lot and started training at a young age”, even though this answer choice was meant to represent an individual who does a lot of physical activity. The vague wording prevents us from being able to use these questions as attention checks.

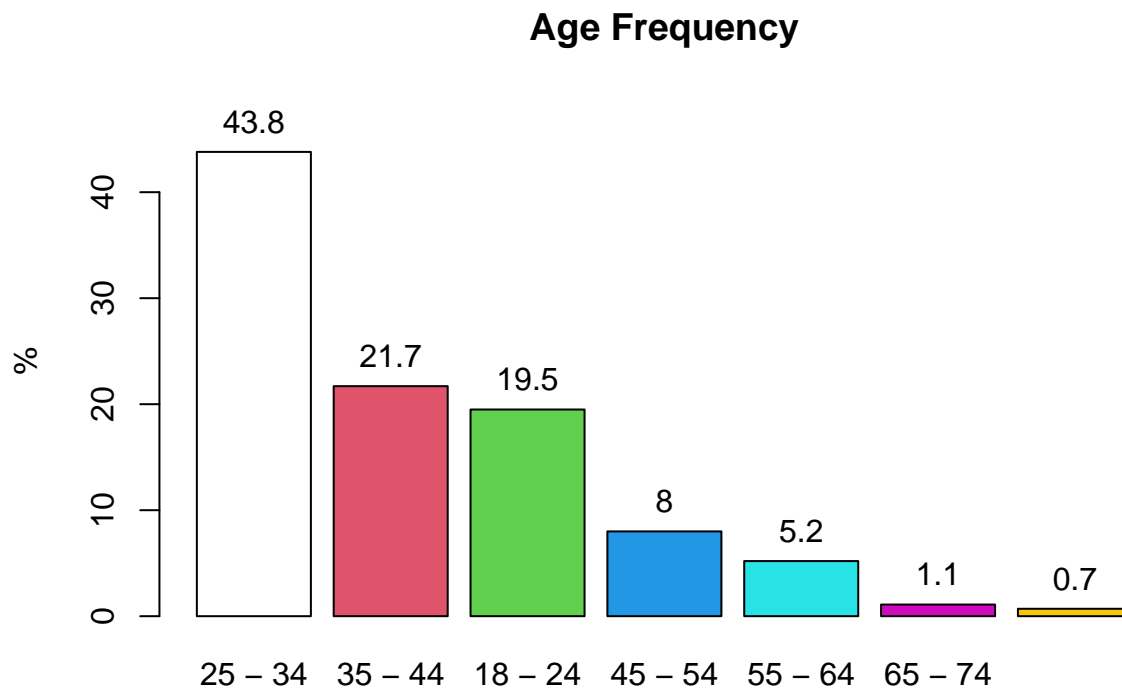
Spillover

Because we deployed this experiment using LinkedIn hashtags, there’s a possibility of spillover if an individual comments on the post. After reviewing the posts, only 1 post had comments and they did not discuss the survey. Therefore we believe spillover to be negligible.

EDA

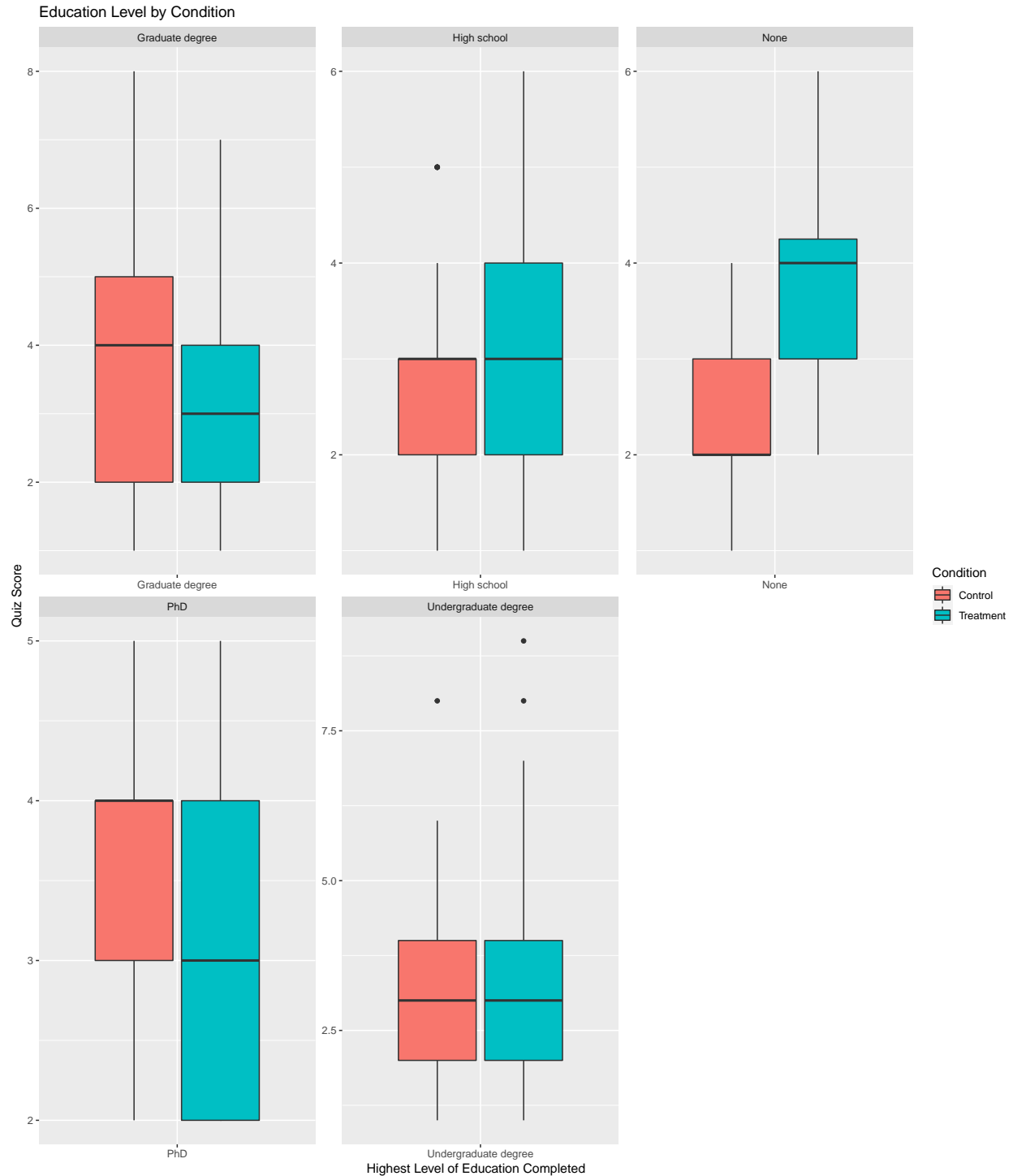
Before we report the regression results, we first examine some of the covariates. This experiment has participants from 33 countries around the world, however, the majority of participants came from the United States so we did not find this to be a useful covariate.

We found that most people were between the ages of 25-34:

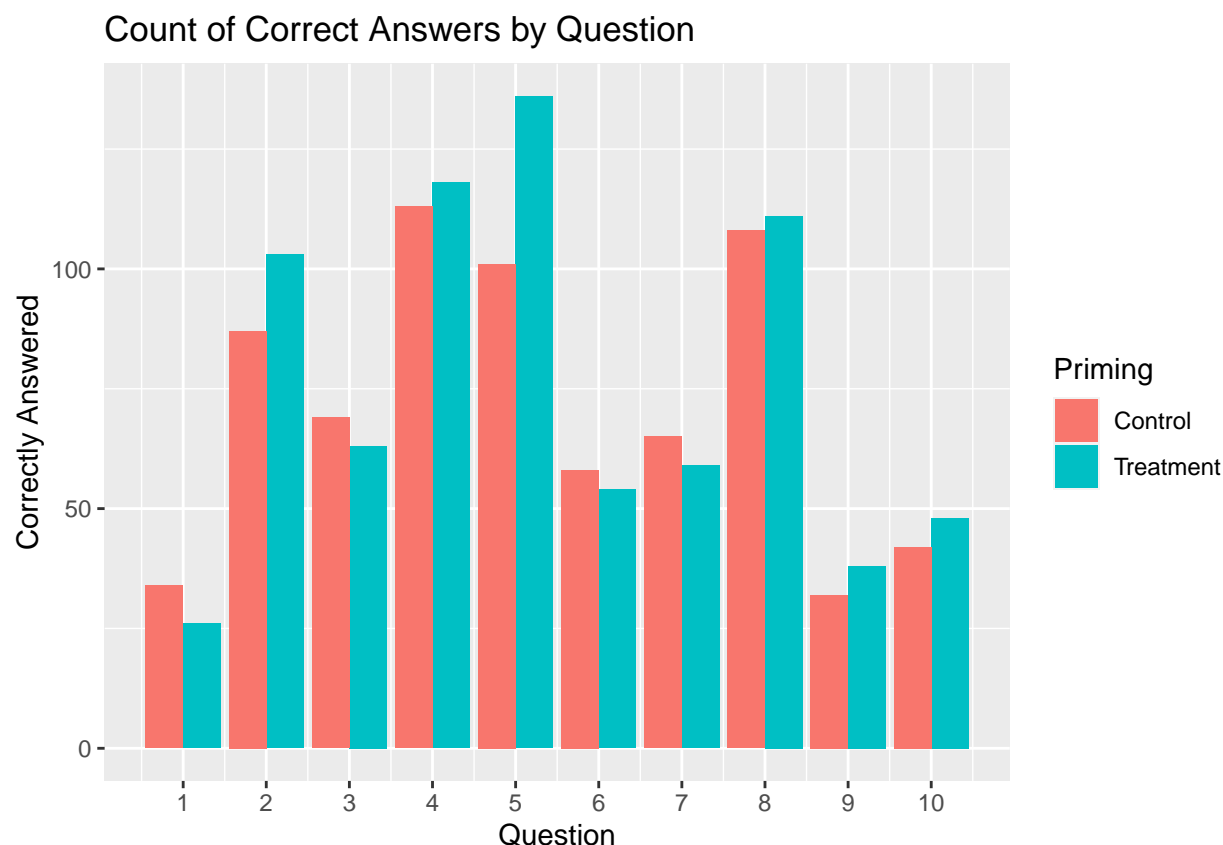


```
## d_without_attrition$age :  
##      Frequency Percent Cum. percent  
## 25 - 34      202    43.8         43.8  
## 35 - 44      100    21.7         65.5  
## 18 - 24       90    19.5         85.0  
## 45 - 54       37     8.0         93.1  
## 55 - 64       24     5.2         98.3  
## 65 - 74        5     1.1         99.3  
## Under 18       3     0.7        100.0  
## Total       461   100.0        100.0
```

Initially, we were motivated to look for heterogeneous effects in regards to education because we were curious if those who have higher levels of education were less susceptible to a subliminal priming message. In this experiment, we hypothesized that there would be a positive relationship between being exposed to the subliminal priming message and the quiz test score. We believed this even for participants with higher education levels. What we did not expect to see was that participants with a graduate degree or PhD would appear to do worse on the quiz when exposed to treatment, as shown in the graph below. This motivated us to look for heterogeneous effects in our regression models.



An area of concern we have is that the subliminal message may only have an effect for a short period of time and that a 10 question quiz might be too long to accurately identify if an effect did occur. To investigate this, we created the graph below which shows no obvious systematic pattern of differences between control and treatment correct answers as people continued through the quiz. Instead, the differences observed appear to be mostly random. Additionally, we ran a regression (reported in Table 4) on just the scores achieved on the first quiz question to see if there was an effect of subliminal messaging on the most immediate question participants answered after receiving the control/treatment message.



Results

For model 1, we started off looking at the effect of the subliminal priming message on the score obtained by participants on the survey quiz, along with a blocking indicator identifying if individuals were *novices* (0) versus *experts* (1) based on their years of experience in the field of data science. We created these two blocks because we believed that *experts* were more likely to score higher than *novices*. However, accounting for this difference in experience, our model still does not show any statistically significant ITT.

For model 2, we added a covariate indicating if participants took the help of external resources. Participants leveraging external resources on the quiz is a neutral control here and controlling for it should help to reduce the variation in the quiz score outcome. That is what we see in the table result as well.

For model 3, we also wanted to control for the age of the participant. Younger in age is noted in this model as an indicator covariate for participants under the age of 24 versus those above that. We see that our standard error reduced by 0.001.

For model 4, we added in the participant's education level. Controlling for education did not really change the point estimate and the standard error stayed roughly the same.

In all of these four models, we do not see any statistically significant effect of the subliminal message. The confidence interval range for model 4 is -0.44 to 0.068, showing an ATE estimate close to zero. Additionally, among all of the models with additional covariates, we find that the *Resources Used* indicator is statistically significant, and perhaps surprisingly has a negative coefficient around -0.6 to -0.7, indicating that those participants who used external materials to try to find answers to the questions achieved lower scores overall. One interpretation for this could be that those who knew less about data science may be more likely to consult external resources to get the answers, but they may not have the skills or intuition to parse the relevant data and find the correct answer.

We also find that all the longer models have a statistically significant negative coefficient for the *Experience* indicator, meaning that those classified as *experts* tended to score lower than those classified as *novices*. We don't have an easy intuitive theory to explain this observation. One speculative idea is that perhaps people who genuinely have more expertise in data science would take the survey less seriously or put less effort into answering it, since they probably make more money and therefore are less motivated by the financial incentive.

Finally, even though we did not find an effect, we did plan to look at heterogeneous effects in regards to education, the results of which are in Table 3.

Table 2:

	<i>Dependent variable:</i>			
	Number of Questions Answered Correctly			
	(1)	(2)	(3)	(4)
Subliminal	−0.181 (0.133)	−0.173 (0.131)	−0.180 (0.130)	−0.188 (0.131)
Experience	−0.158 (0.167)	−0.317* (0.164)	−0.408** (0.165)	−0.519*** (0.174)
Resources Used		−0.610*** (0.158)	−0.676*** (0.160)	−0.651*** (0.163)
Younger In Age			−0.389** (0.167)	−0.277 (0.172)
HS Diploma				−0.429 (0.358)
Undergrad Degree				−0.200 (0.346)
Masters Degree				0.076 (0.361)
PhD				0.110 (0.363)
Constant	3.289*** (0.102)	3.749*** (0.158)	3.898*** (0.169)	4.023*** (0.369)
Observations	464	461	461	461
R ²	0.006	0.044	0.055	0.069
Adjusted R ²	0.002	0.038	0.047	0.053
Residual Std. Error	1.437 (df = 461)	1.400 (df = 457)	1.393 (df = 456)	1.389 (df = 452)
F Statistic	1.390 (df = 2; 461)	7.014*** (df = 3; 457)	6.644*** (df = 4; 456)	4.192*** (df = 8; 452)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3:

	<i>Dependent variable:</i>				
	No HS Diploma (1)	HS Diploma (2)	Number of Questions Answered Correctly Undergrad Degree (3)	Graduate Degree (4)	PhD (5)
Subliminal	0.875 (0.634)	-0.083 (0.284)	-0.118 (0.202)	-0.576* (0.312)	-0.232 (0.279)
Experience	-1.000** (0.413)	-1.023*** (0.390)	-0.607** (0.307)	-0.241 (0.345)	-0.657** (0.290)
Younger in Age	-1.500** (0.584)	0.097 (0.308)	-0.319 (0.235)	-0.303 (0.529)	
Resources Used		-0.127 (0.558)	-0.655*** (0.226)	-0.769** (0.372)	-0.339 (0.341)
Constant	3.000*** (0.413)	2.972*** (0.598)	3.815*** (0.225)	4.299*** (0.402)	3.986*** (0.380)
Observations	15	78	213	107	48
R ²	0.452	0.047	0.058	0.095	0.117
Adjusted R ²	0.302	-0.005	0.039	0.060	0.057
Residual Std. Error	1.103 (df = 11)	1.248 (df = 73)	1.481 (df = 208)	1.478 (df = 102)	0.976 (df = 44)
F Statistic	3.022* (df = 3; 11)	0.908 (df = 4; 73)	3.173** (df = 4; 208)	2.687** (df = 4; 102)	1.941 (df = 3; 44)

Note: * p<0.1; ** p<0.05; *** p<0.01

In Table 3, we looked at the number of questions answered correctly relative to a participant's level of education. We see that subliminal priming tends to have a negative impact on participants with a graduate degree. Additionally, participants who did not attain a high school diploma, have a confidence interval of -0.36 to 2.11, showing a potentially positive relationship. However, the low number of observations for these two groups makes us cautious about making any strong claims about the validity of these results.

Out of curiosity, we wanted to see whether subliminal effects tended to just have a shorter time span of an effect, and so we created a model looking at the treatment effect based on the answer to only the first quiz question. However, we do not see anything noteworthy here to report. The hypothesis that subliminal stimuli have a short-term effect on the performance of individuals on a quiz cannot be meaningfully measured here and needs further research.

Table 4:

<i>Dependent variable:</i>	
Question 1 Correctly Answered	
Subliminal	-0.051 (0.634)
Experience	-0.003 (0.413)
Younger in Age	-0.061 (0.584)
Resources Used	-0.046
Constant	0.199 (0.413)
Observations	466
R ²	0.013
Adjusted R ²	0.005
Residual Std. Error	0.332 (df = 461)
F Statistic	1.573 (df = 4; 461)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Conclusion

In this study, we did not find a statistically significant impact of subliminal priming messages on quiz scores. Also, since this study was conducted in regards to data science concepts we cannot generalize this to other specialties. Given the potential biases, assumptions and data quality issues, we recommend future experiments be conducted, and we discuss potential issues and improvements here as a guide to possible further research.

Limitations and Future Enhancements

While we had a survey response with 518 participants, the overall results were not compelling evidence of any subliminal effect nor did the user participation quite pan out as we had hoped. From a perspective of carrying out an online survey, we noticed some underlying limitations in the way Qualtrics can differentiate between good and bad survey responses. While Qualtrics notes that it can identify web bot responses, it was not always able to catch bad responses. Given that randomization was in play, the web bot responses were considered to be equally distributed and not considered as consequential in nature, though they were a knock on the statistical power, which meant that having more responses would have helped.

Another possible effect that could have polluted the outcome was the gift card incentive that may have provided a different motivation to people taking the quiz, and the presence of responses in the data from people who seemingly answered the quiz multiple times to maximize their chances of winning a gift card seems to confirm this was an issue. This brought us to question whether it would have made more sense to conduct the experiment without incentivizing it, though this could have reduced our chance of getting a decent number of responses. It likely would have necessitated running the survey for a much longer period of time.

Overall, while the survey quiz was geared at individuals working in the field of data science, its generalizability was limited since assessing people in a certain field of their expertise does not always translate into a quiz based assessment. Subliminal priming can be carried out in many different ways, some that are more impactful than others, as in the case of visual priming that could provide emotional stimuli that are known to improve memory processing. As such, generalizing the experiment to different areas of expertise may require developing different methods for administering a priming stimulus and of measuring participants' performance.

If we were to look at a future iteration on this experiment, it likely would not be carried out through an online survey. Testing for an effect through subliminal priming needs a more robust mechanism of measure. It could be that an improvement in a person's ability could be measured in an area where they feel at ease in working but feel underrated or feel that they could do better. In such cases, a difference in difference measure would help in determining how well the person does on a task they are comfortable with, followed by measuring their effectiveness on a similar task after a subliminal priming treatment. This removes the problem that people are not always effective at something new they are asked to put their mind to, and removes anxiety-related variances in responses.

Another aspect to consider here is that cognitive studies on subliminal priming point to short experimental observation periods. While the number of questions asked on the quiz were in line with the statistical power calculations for determining a difference, the quiz format did not lend itself to a short observational period. Avoiding a ceiling effect in scores meant that the quiz needed to be a little more challenging to participants and not too easy for the treatment and control groups to breeze through. This may have increased the length of time participants took to complete the questions, which may have reduced the measurability of any priming effects.

Finally, the whole premise of the experiment can be completely thrown off if there is no good way to measure if subliminal priming even took place in an individual. Event-related potentials (ERPs) are brain signals that arise as the result of a thought, internal stimuli, or an individual's perception of an external stimulus. ERP signals can be measured through EEG (Electroencephalography) signal measuring instruments and

in an ideal setting would potentially allow a way for this experiment to definitively identify compliers to subliminal priming prior to measuring the effect it had on the individual.

Appendix

Treatment

If assigned to treatment, participants received several sentences motivating them to imagine being an expert. If assigned to control, they received several sentences about being a journalist in today's technology-driven environment.

The treatment is as follows:

“Data Scientist is the sexiest job of the 21st century! Now imagine you are the most talented and motivated Data Scientist in the field. You are very expensive to hire and, given the very competitive market for your services, it is also difficult to retain you. You are one of the few people with a strong combination of scientific background, computational and analytical skills.”

The control is as follows:

“In the world of dining, digital photography and platforms like Instagram have become the main method that restaurants use to communicate with patrons. Rocket launches are now live-streamed online, which lets space reporters watch from their phones instead of heading to the space station. And in the entertainment world, video streaming has opened doors to a wealth of new content — so much that reporting on movies and TV shows has become an art of curation.”

Resources

- Self-affirmation improves performance on tasks related to executive functioning
 - <https://www.sciencedirect.com/science/article/pii/S0022103116302840>
- Long-term Effects of Subliminal priming on Academic Performance
 - <https://www.gsb.stanford.edu/faculty-research/working-papers/long-term-effects-subliminal-priming-academic-performance>
- Cognitive consequences of affirming the self: The relationship between self-affirmation and object construal
 - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3149801>
- Subliminal Priming—State of the Art and Future Perspectives
 - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6027235>

Qualtrics Survey Flow

Survey flow

Published

