Expectations in Unfamiliar Territory: How Reinforcement Impacts Using an Unfamiliar Device

Adam Sebetich, Heng Yon, Kevin Ke, Joel Elizaga, Tyler Bauer
The Information School
University of Washington
Seattle, WA USA 98195
{sebet86, heng008, kevindke, jelizaga, tbauer16}@uw.edu

ABSTRACT

This project involves discovering the psychological aspect of human behavior with relations to an uncommon digital device. The goal is to understand if giving participants a good or bad review of a product will have an affect on their thoughts and behavior about themselves. This is highlighting the fascination, mystery and similarity of the placebo and nocebo effects. Specifically, it is highlighting how positive and negative reinforcement can affect different research experiments in the scientific community. Those effects will be targeted in this research experiment, which highlights the usage of an unfamiliar mouse. In this case, the experiment involved the usage of a trackball mouse to see the participant's performance on a clicking activity. After the task ended, participants were asked to rate themselves on what they believed their performance was on a scale of one to ten, (one being bad and ten being great). Our experiment indicated that letting our participant know that previous participants viewed the task as easy or difficult did not have a significant outcome on their own perception of their abilities.

Classification Keywords

Placebo, positive reinforcement, negative reinforcement, mouse, unfamiliar.

INTRODUCTION

Believing is one of the most powerful abilities we have as humans. When we truly believe in something such as religion, people or products, it can completely alter the way we think about ourselves. For instance, "the mere act of taking a pill may elicit" one to believe they are "improving"[3] their own health. In parallel by having a strong understanding and belief in something, we can change the way we view how we perform using an unfamiliar object. Having a preconceived idea of an unfamiliar technology can also lead to us potentially viewing our usability as positive or negative regardless of how we actually performed. This idea can affect the technological community and could lead to a reshaping of the way they advertise various unfamiliar new products to potential users. This experiment monitored participants and gave both positive and negative reinforcement to participants while using an unfamiliar device to see if it would have an effect on their own perception while performing a certain task.

RELATED WORK

There have been several studies and papers written about similar topics to ours, encompassing either individual pieces of our paper or individual pieces related to our paper, yet none are as scoped down and as focused as to look at reinforcement on using a trackball pointing device.

A paper from 1986 performs an experiment to test whether positive or negative reinforcement can make an impact on "performance as a function of extraversion-introversion" [1], meaning they have a between subjects factor of extraversion-introversion included in their test. They used two factors, the second being the type of reinforcement, to determine if there was any type of interaction between the two factors. The tasks the participants were asked to perform were a computer game, where they used the cursor to find a hidden target, and some complex arithmetic in the base-6 number system. They found that the interaction between the two factors was significant, and that extraverts did better with positive reinforcement and introverts did better with negative. While these results are fascinating and significant,

they dwell outside of our scope and what it is we are looking for.

Another paper from 2005 looks at how "the ERN predicts the degree to which participants are biased to learn more from their mistakes" [4]. This paper focusses mainly on how people learn from their mistakes, and what impacts how much is learnt from these mistakes. Specifically, a set of personalities called positive learners learn more from not making mistakes, while the negative learners tended to learn more from actually making mistakes. This paper has some overlap with our experiment, as it deals with positive and negative reinforcement. Here, negative reinforcement is said to lead to a greater impact on learning, as one learns from their mistakes. This contradicts our hypothesis, as we predict that people who have a negative emotion to a task will perform worse. This paper provides a direct contradiction to our experiment by looking at repeated tasks with feedback, rather than an initial task where the participant is primed.

METHOD OUTLINE FOR PARTICIPANTS

When deciding who would be our participants for this experiment, we limited the scope to individuals who are in the University of Washington Information School. The individuals who were sought out were chosen at random and were participants who were willing to help out of their own generosity. In total, 20 participants were chosen because of the time restrictions of the research experiment being roughly 6 weeks long. The participants were chosen on two separate days to try and widen the pool of potential participants as best as we could. After asking the participants if they were willing to help in the experiment, they were guided into a separate room and read a standardized informational script by the proctor of the test.

The iSchool students that participated in the research experiment were told the same script up until the portion of the script where the variable of either negative or positive reinforcement was given. The same proctor was used throughout the entirety of the experiment to avoid outside variables from affecting the results, as well as to project the same "'aura' of competence and trust for the patient"[5]. The proctor would notify the participants that the all the previous students completed the test with either extreme ease, or extreme difficulty. This variable was also chosen at random to eliminate potential bias from the proctor. Thus, introducing a reinforcement effect that might affect how the participants take the test. After that last sentence, the proctor would allow the participant to initiate the experiment potentially leaving them with the positive or negative reinforcement as the last thing the participants hear.

The experiment itself involved the participants utilizing an unfamiliar device, the trackball mouse. We were able to measure if the trackball mouse was considered unfamiliar to the participants by asking them if they were familiar with the device before continuing with the test. If the participants stated they were above a 2 on a scale of 1-10, 1 being unfamiliar and ten being a professional user, then they were not allowed to continue with the experiment.

METHOD OUTLINE FOR CLICK TEST

After hearing the script, the participants initiated the point and click test. The point and click test was developed by one of our team members and involved a total of 20 small black circles appearing in sequence on the screen. 20 circles was chosen because that allows the test to be both long enough to engage the participants, but also short enough to allow us to conduct multiple tests in a short period of time. The participants would first click a small black circle in the middle of the screen to initiate the test.

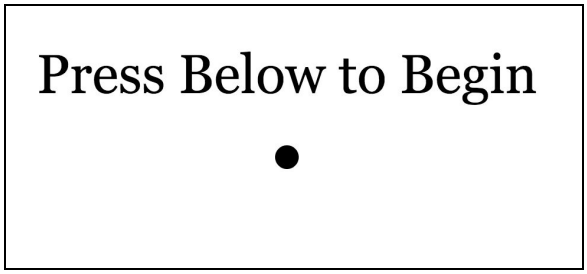

Figure 1.
Each of the black circles that appeared on the screen were always in the same order to keep the test consistent for each participant.

Clicking a black circle would immediately prompt the next circle to spawn, Figure 2. The measure of speed between clicks, error rates, and degree of errors (if errors were made) were all recorded by our developed application. Upon completion of the test, a one question survey asked the participant to rate how well they believed they performed the point and clicking test, again on a scale of 1-10, 1 being bad and 10 being great, Figure 2. The test concluded after the
participant submitted their response and a standardized script for the ending of the experiment notifying the participant of our intent was read.

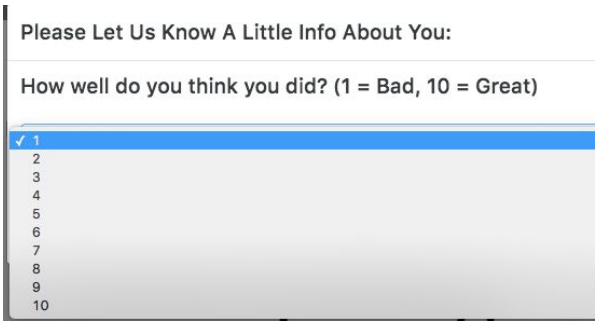

Figure 2. Post test survey for subjects to enter in

IMPLEMENTATION DETAIL

The point and click test utilized a total of 114 HTML lines of code and 277 Javascript lines of code. The coding languages used in the development of our application were primarily HTML and Javascript. A website was developed for the point and click test and all recorded information was stored on Firebase. Firebase allowed us to gather the survey responses and the other measured variables as well and stored them in the Cloud for easy retrieval from anywhere.

MEASURES AND HYPOTHESIS

Five measures were collected throughout each participant's test: (1) initial familiarity (2) speed between clicks, (3) misclick/error rate, (4) degree of errors, and (5) post test self-performance rating. The primer measure of initial familiarity allowed us to determine if the participant's level of knowledge of the device would be in our range of unfamiliarity, capping off the eligible mark at a max rating of 2.

Hypotheses were that, for the participants who received the negative reinforcement would rate themselves lower than those who received the positive reinforcement.

RESULTS AND DISCUSSION

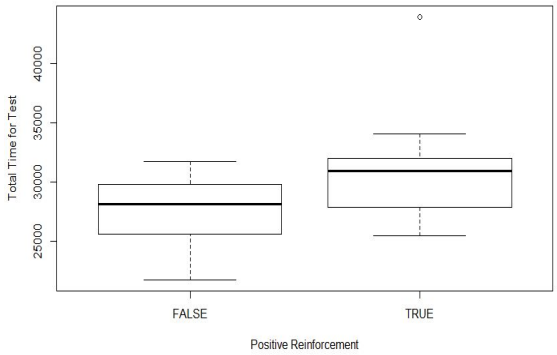

Figure 3. Boxplot of the data to see if there was a difference between the amount of time taken to complete the test for each group.

Our results of our first test concluded with some unexpected results. Our initial t-test to determine if there was a difference between the amount of time it took to complete the entire point and click test between the two groups (receiving positive reinforcement vs receiving negative reinforcement) concluded with a marginal trend. The type of reinforcement (positive or negative) received did not have a significant impact on the speed at which a trackball pointing device was used ($t(14.751) = 2.0382$, $p = 0.06$), although the trend suggests that positive reinforcement helps and the positive "placebo [may have] reduc[ed] anxiety or other negative emotions"[2] our subjects may have had prior to taking on the task. A more definitive conclusion could be made with further research. With a difference that does not support any difference, there is no reason to pursue any further tests with this same dataset.

With a p-value slightly above 0.05, we were very close to determining how people's perceptions of how difficult a task is may determine their ability to do said task. The marginal result, while not

significant, is interesting in the sense that it only raises more questions than answers. Our test concludes with a nonsignificant result, but further testing may sway either direction. Our procedure may have had an impact on our results, as well as our relatively small sample size. A larger scale experiment in a more controlled environment may result in findings that are slightly more significant than ours, confirming our hypothesis. Or the findings may swing the other direction and provide results that drastically put our hypothesis to shame. Either way, our preliminary testing concluded that there is at least some plausibility in our hypothesis, which was a stretch to even begin with.

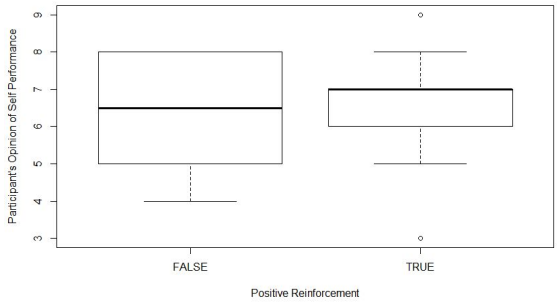

Figure 4. Boxplot of the participant's self assessment of their performance for each group.

Our secondary test we performed was a t-test to see how the type of reinforcement may have impacted a user's satisfaction or confidence in using an unfamiliar device. The results of this test were not as promising as the previous. The type of reinforcement received did not have a significant impact on the user's own opinion about their performance ($t(17.653) = 0.29002$, $p = 0.7752$). With results as nonsignificant as this one, there is not much of a chance that any further research or larger sample sizes will find any contradictory results. Due to the results having such a high p-value, there is no reason to do any further testing with this dataset.

With a p-value of 0.7, we can be sure that there is not much difference between two groups in terms of their own self assessment. Looking at Figure 3, there is an obvious overlap between the two groups, with the positive reinforcement group almost entirely included within the other group's first and third quartiles. It is safe to say that what you tell a person about something has little to do with how they feel about using it, relative to other people's performance. Most people, regardless of the group they were in, rated themselves at a 6 or 7 in their self assessment.

CONCLUSION AND FUTURE WORK

One's perception of how difficult or easy an upcoming task will be, told by an external source, does not psychologically affect their personal rating after completing the task. While telling a person an unfamiliar device may slightly make them perform a task faster, we cannot determine if the difference is large enough for it to make a difference, at least with a device they are unfamiliar with. Overall, regardless of the reinforcement a person is given, they will perform an unfamiliar task with the same outcome.

Although the data collected did not directly support our initial hypothesis, we plan to pursue further tests with larger sample sizes to determine to what degree does a positive reinforcement impact one's performance. Such feedback will give a stronger grasp behind the psychology of performing a task, as well as directly applicable benefits like training professionals.

ACKNOWLEDGMENTS

The authors and researchers would like to thank Paul Weiss, Jacob Wobbrock, and the entire iSchool community for their support.

REFERENCES

1. Boddy, John, Carver, Annabel, and Kevin Rowley. 1986. Effects of positive and negative verbal reinforcement on performance as a function of extraversion-introversion: Some tests of Gray's theory. *Personality and Individual Differences*. Volume 7, Issue 1. Pages 81-88
2. Flaten, Magne A, Aslaksen, Per M., Lyby, Peter S., and Bjorkedal, Espen. 2011. The Relation of Emotions to Placebo Responses.
3. Foerster, Elisabeth. The Placebo Effect. 2016. Retrieved December 10th, 2016 from

http://www.immpressmagazine.com/the-placebo-effect/

4. Frank, Michael J., Woroch, Brion S., and Tim Curran. 2005. Error-Related Negativity Predicts Reinforcement Learning and Conflict Biases. *Neuron*. Volume 47, Issue 4. Pages 495-501
5. Nasarallah, Henry. 2011. The Most Powerful Placebo is Not a Pill.Retrieved December 11, 2016 from http://www.mdedge.com/currentpsychiatry/article/64401/most-powerful-placebo-not-pill