Can Setting An Expectation Affect Test Performance?

Examining the effects of setting an expectation of test difficulty on subject’s performance and levels of stress and arousal.

By Alex Thele

When I first was tasked with designing an experiment, I set out to observe the effect that setting an expectation of test difficulty has on performance. This idea came from past personal experience and anecdotes from other students about performing poorer on tests that we were told would be hard regardless of their actual difficulty level. I hypothesized that a task subjects were told was easier would take them less time and be completed with more accuracy than a task that they were told was hard.

Experimental Design

In order to test this hypothesis, an experiment was designed with two “treatments”, or different conditions: Easy and Hard. These tests were pattern-matching (Figure 1) tests of the same difficulty level, with the only difference between them being how difficult the subjects are told they would be. Before each of the tests were given, the subject was told the average time it took other subjects and the level of accuracy they should be able to get: 40 seconds with 100% accuracy for Easy and 1 min with 90% accuracy for Hard. These numbers were fictitious and told in order to further set up the supposed difference in difficulty between the tests. Additionally, “HARD” or “EASY” was written on the top of each exam in large, bold letters to further emphasize the difference.

One of the most important things when designing an experiment is to make sure that you eliminate as many confounding variables as possible and make sure that the only thing that changes is the independent variables that you are manipulating- in this experiment, those were the two treatments or supposed difficulty levels of the tests. There were four total tests, and since this experiment used a within-subject design, each subject received both
treatments. This means each subject took two Hard tests and two Easy tests. The tests are given in alternating difficulty level, and in order to rule out the testing order as an influencing factor, subjects are placed into one of two groups: group 1 (Hard1, Easy1, Hard2, Easy2) or group 2 (Easy1, Hard1, Easy2, Hard2). When statistical analyses were run, there was no significant difference between the groups, which means that the order that the tests were given did not affect the subject’s performance or stress level.

Both behavioral and physiological data was collected. The behavioral data collected was straightforward, consisting of just recording the time it took subject’s to complete each test and speed and determining what percent of their responses were correct to determine their accuracy. The physiological data we measured were galvanic skin response (GSR) and heart rate, both of which are controlled by the autonomic nervous system, which is what controls everything that we do not think about consciously. Consequently, these are two of the main measurements involved in a polygraph test, which additionally measures respiratory rate (Pflanzer). Both heart rate and GSR are measurements of the subject’s arousal and stress level, with GSR being basically a measure of sweat production. These values are measured using five electrodes, one of each of the following (figure 2): inside of left ankle, inside of right ankle, middle finger, index finger, inside of wrist, (with the finger and wrist electrodes on the non-dominant hand). GSR and HR values measured during the testing are compared to baseline values from non-stressful conditions; in this case, while watching a clip from YouTube entitled “The World’s Most Relaxing Film”.

![Figure 2. Placement of the electrodes used to measure the physiological variables (GSR and heart rate). (Pflanzer)](image)
**Effects of Expectation**

As can be seen in Figure 4, there was a statistically significant difference between the amount of time it took for participants to take the Easy tests (51.4 seconds) and the Hard tests (58.7 seconds). However, while there was a difference between the subjects’ accuracy between the Easy tests (84.7%) and the Hard tests (61.1%), there was not statistical significance due to the high standard errors (Figure 3). Standard error shows how spread out data is on either side of the mean value reported, and the larger error bars on this graph mean there is more variation to the data and it caused the difference between the accuracy to be not significant.

![Test Speed](image1)

**Test Speed**

![Test Accuracy](image2)

**Test Accuracy**

One of the scientific contexts these results can be placed in is a psychological phenomenon called the Pygmalion effect, a sort of self-fulfilling prophecy regarding performance and expectation.

*The Pygmalion effect*

The Pygmalion effect is a psychological phenomenon that can be seen in education, as well as the workplace and other environments with leader-learner dynamics (Figure 5). This was documented in the 1960’s by psychologist Robert Rosenthal in a series of experiments where he went to elementary schools and gave students an IQ test, after which he gave teachers a list of students who had the potential to succeed; these selected students excelled...
academically compared to their peers. However, the list he gave the teachers was generated randomly and independently of the test results, which the increase in student’s success was not due to higher IQs, but to the higher expectations placed on them by the teacher (Duquense University). The effect can be summarized in this quote from the researchers themselves: “When we expect certain behaviors of others, we are likely to act in ways that make the expected behavior more likely to occur.” (Rosenthal and Babad, 1985)

**Relationship to Test Anxiety**

While my main hypothesis focused on the effect of expecting a certain difficulty level on test, I was also curious if there was a correlation between test anxiety and performance. After finishing the testing for each subject, I asked the participant to rate their general level of anxiety before a test on a scale of 0 to 5, with 5 being the most anxious. The self-reported responses varied across the range, but the majority of subjects reported a test anxiety level in the middle; the mean, median, and mode of the responses were 2.66, 3, and 3, respectively.

There was no significant correlation between any of the performance measurements (time, accuracy) or the GSR data for either of the testing conditions. However, there was a significant correlation between heart rate and test anxiety for both the Easy and Hard condition (Figure 6,7). There was a negative correlation between the subject’s level of anxiety and their change in heart rate for both conditions; that is, as their level of test anxiety increased, the difference between their heart rate and baseline decreased. The $R^2$ values for these figures indicate a relatively strong correlation. This is due to the subjects with test anxiety having an elevated heart rate throughout the experiment, regardless of whether they were being tested or being show the video.
These findings are interesting, and certainly add credibility to the subjects’ self-reported ratings of their test anxiety level. It may be possible that subjects who rate themselves high on the test anxiety scale have a higher baseline level of anxiety in general, and possibly have some form of generalized anxiety disorder. Regardless of these implications, one thing is clear: setting reasonable expectations about a test’s difficulty level and encouraging students will lower their stress and improve their performance.
Works Cited
