COGNITIVE PERFORMANCE & ECONOMIC INCENTIVES

BY
Morganne Beck, Carola Clift, and Michael Ensley

Introduction

On any given day, the psychology department of any university is running countless psychology experiments. Often, university psychology labs don't get tremendous funding to pay their subjects nor do students just volunteer to be subjects. Accordingly, laboratories will find other ways to pull in participants. At the University of California, San Diego, students taking lower division psychology courses are required to participate in three hours of experiments. Should the students not do so, their grade will be lowered. Can demanding students to participate in experiments of which they have no knowledge of purpose, no desire, and no invested interest affect the students' performance? On the other hand, can paying subjects money affect performance? The questions at hand have different answers depending if you follow the psychologists' approach to experimentation or the economists' and political scientists' approach to experimentation.

As a psychologist, you might claim that regardless of how a participant happened upon an experiment, experimenters can assume the subject will inherently "do their best." Economists and political scientists, for the most part, think however, that people will perform to the best of their ability only when given an incentive to do so. It follows that while the participant who is not getting paid and must be there, will perform well enough. Meanwhile, the participant who is getting paid will have an invested interest and perform better.

This phenomenon can be explained by the induced-value theory. The main idea behind the induced-value theory is that a reward medium can be used to "induce prespecified characteristics in experimental subjects, and the subjects' innate characteristics become largely irrelevant." Unfortunately, class credit and untampered grades aren't considered a reward medium; money is however. A subject's innate characteristic, when not induced to perform a certain way through monetary compensation, might be to finish as fast as possible. There is no guarantee these subjects will commit to the concentration and effort required of most experiments. Vindictive subjects might even hinder experiment results with aims of sabotage.

Is the only way to get quality results, then, to pay all your subjects? Why wouldn't a subject show up just for the money and still not try their hardest? The induced-value theory further accounts for these queries by laying out three conditions necessary to suffice the prespecified characteristics experimenters strive to achieve. These three conditions can be described as follows:

1. Monotonicity: With monotonicity, the subject is enticed by more of the reward medium, and never less, without becoming satiated. For example, more money is always better than less. But if the reward medium were chocolate chip cookies, a person will want more to a certain point, and beyond that point, more cookies are undesirable (whether due to a full stomach, unwanted calories, or fear of cavities!).
2. Salience: Salience is achieved when a subject is rewarded directly based on their actions or performance. Paying a subject a $5.00 flat rate for showing up is not salient because what the subject does once the experiment has started is unrelated to the reward medium. If, however, a subject is given points throughout the experiment for certain actions, salience has been achieved through the incentive to earn more points.
3. Dominance: Dominance states that it is solely the reward medium, and not some other factor or innate characteristic of the subject, which is responsible for the prespecified characteristics the experimenter wished to test.

The Experiment

A Classic Psychology Experiment
Since the aim of our research and experimentation was to show the effectiveness of the induced-value theory, we first had to find an experiment which we could manipulate in an attempt to show that monetary incentives can induce subjects to perform better. We chose Craik and Tulving's experiment on the effect of encoding on memory recall specifically because it is a classic psychology experiment that has been studied in thousands of classrooms since their results were published. Craik and Tulving's hypothesis was that it is the qualitative nature of the operations carried out on the material, not the intention to learn, that determines memory recall. They claimed that a subject's ability to correctly remember a word corresponds to the depth of processing he or she performs at the time the material is presented. The more thoughtful the analysis of the word, the deeper the processing; and the deeper the processing, the better the recall. For example, if a person is only required to determine the color a word is printed in, he will only think about the word long enough to say that it is written in red. Such sensory perception only requires analysis of the color of the ink and not the actual meaning or structure of the word. Consequently, only a weak memory trace is created. On the other hand, if a person is asked to decide whether or not a word fits into a particular context, analysis of the word's meaning will result in a stronger memory trace. Craik and Tulving hypothesized that a subject's ability to correctly remember a word corresponds to the way in which the subject processes the word at the time it is presented, with deeper processing yielding better recall.

To further support their hypothesis that the operations performed on the material determine retention, Craik and Tulving established trials under both incidental and intentional learning conditions. Incidental learning refers to learning that is a product only of the encoding operations performed on the material. An example of this is your haphazard observation that the man in front of you at the post office has blond hair. Intentional learning, by contrast, would result from a concerted desire, on the part of a subject, to learn material that he or she expects to need later. If you suspect the man in front of you is about to take your wallet, you will actively note his hair color, height, and more, in order to describe him to the police later. Craik and Tulving argue that a subject's intention to learn does not affect his or her ability to remember.

To test their hypothesis, Craik and Tulving ran a series of experiments requiring subjects to process sixty words to different depths by asking them one of three types of yes-no questions about a common noun. The experiments were performed on two groups of students in a classroom setting. The three types of questions presented to the subjects, each inducing deeper processing respectively, were: (1) if the word was typed in capital letters, a question requiring little thought, or (2) whether the word rhymed with another word, requiring more thought but little analysis, or (3) if the word fit into a given category, requiring the most analytical thought about the meaning of the word. After the subjects were asked one of these questions, the word was visually projected upon a screen for exactly one second. The subjects then had exactly five seconds to answer

the yes-no question. This was repeated with sixty common nouns; each type of question was asked 20 times, with half yielding "yes" responses, and half yielding "no" responses. The order of the questions was random. After all sixty words had been presented, the subjects were given a quiz in which the subjects were told to pick the 60 words they had seen from a list of 180 common nouns.

Craik and Tulving ran trials in which they created both intentional learning and incidental learning environments. In some of the experiments, the subjects were not told that there would be a quiz administered until after the words had already been presented. This was an incidental learning environment. In other trials, they informed their subjects before the experiment began that they would be quizzed on the words later. This created an intentional learning environment. In the intentional learning environment, subjects were expected to use a memorization technique (i.e. saying the word out loud or repeatedly to themselves) when presented with the 60 words in order to perform better on the quiz. This proved not to be the case.

Craik and Tulving's results for both groups showed that recall was best for those words that had been processed most deeply. Table 1 shows that the case response type questions, which required the least amount of thought to answer, yielded the smallest proportion of words remembered correctly. The rhyme response type questions yielded slightly higher recognition. The category response type questions, however, yielded the highest proportion of words recognized correctly. The hierarchy in their results supported their hypothesis, which we accept: the nature of the operations performed on material at the time it is presented does affect memory recall. However, since Craik and Tulving found a similar hierarchy in both intentional and incidental learning environments without any significant difference in the proportion of words remembered correctly, they were further able to deduce that intention to learn does not affect memory recall.

Our Design

Contrary to what Craik and Tulving found, we believe the intention to learn does indeed affect memory recall. Based on Craik and Tulving's claims, we formed a null hypothesis: the intention to learn does not affect how well words are recalled. To test our null hypothesis, we introduced the induced-value theory into Craik and Tulving's experiment. We replicated their experiment exactly, except for three differences. First, each word was presented visually to the subject on a computer, rather than being projected on a screen. We used the computers because we were granted exclusive access to the UCSD Political Science computer lab, and because computers can easily be programmed to flash a word for exactly one second. Second, subjects were tested individually because computers lend themselves best to individual use. Third, we introduced the factor of monetary incentives, which we used to induce in subjects the motivation to learn.

To motivate our subjects to learn, we applied the principles of the induced-value theory. We designed our experiment to include an experimental group and a control group. All of our subjects were undergraduates at UCSD, and each trial lasted about 30 minutes. For members of the control group, we replicated Craik and Tulving's intentional learning experiment exactly, except for the use of computers and one-on-one testing, noted above. These subjects were paid a flat rate of $10.00 for their participation in the 30 minute experiment. Members of the experimental group were told before the experiment began that they would be paid 50 cents for every word they recalled correctly during the quiz. Thus, subjects had the opportunity to earn $30.00 in one-half an hour. We assumed that the subjects valued money (the more the better), and our instructions made the relationship explicitly clear between the monetary reward and their performance on the quiz. Designed in this way, our experiment satisfied, respectively, the conditions of monotonicity and salience necessary to suffice the induced-value theory. Our results, discussed below, indicate that the third necessary condition of the induced-value theory, dominance, was also achieved. Achieving dominance allows us to reject our null hypothesis.

Results and Discussion

The results for the control group in our experiment reflect findings similar to Craik and Tulving's. The introduction of induced-value as the independent variable to our experiment, however, has produced results that challenge some of Craik and Tulving's claims about learning. Our results suggest, contrary to the claims of Craik and Tulving, that people can be motivated to learn.

The Control Group

The results from our control group (those who were paid a fixed rate for their participation) resemble Craik and Tulving's experiment in two ways. First, the results for the control group reflect strong encoding effects under intentional learning conditions. In other words, attention to the meaning of the words is a necessary aspect of good retention. Examination of the three response types (i.e. case, rhyme and category) for the subjects in our control group illustrates the effects of encoding (Table 1). Each response type requires increasing levels of analysis to answer. Thus, the words that are analyzed to a greater depth are remembered better. As Table 1 shows, the words that were encoded using a category question were recalled best - 84% recalled correctly. The words encoded using rhyme questions and case questions were recalled less frequently - 65% and 48%, respectively. These results are similar to the Craik and Tulving's experiment, which can also be seen in Table 1. The results for our control group demonstrate that the depth of processing, or the level of semantic involvement, is a useful framework in understanding retention and recall. Thus, our results support Craik and Tulving's claim that attention to the word's meaning is a necessary aspect of good retention.

Table 1
A Comparison of Craik and Tulving and Our Control Group
(Proportion of words recalled correctly)

Response Type	Case	Rhyme	Category
Craik and Tulving	.33	.53	.75
Control Group	.48	.65	.84

Although our results support and agree with the Craik and Tulving hypothesis that recall depends on the qualitative nature of the encoding, it

must be noted that our subjects performed better overall. The performance of the subjects in the control group exceeded the performance of Craik and Tulving's subjects in all three response types. We attribute these significantly higher results to the one-on-one interaction of our experiment, which the Craik and Tulving experiment did not possess.

The subjects in our control group resembled the Craik and Tulving experiment in another important way. The results for the control group produced significant differences for yes and no responses (Table 2), which were also similar to Craik and Tulving's findings.

Table 2
A Comparison of Yes/No Responses
(Proportion of words recalled correctly)

Response Type	Case	Rhyme	Category
Craik and Tulving "YES"	.33	.62	.86
Craik and Tulving "NO"	.33	.42	.64
Our Control Group "YES"	.49	.72	.94
Our Control Group "NO"	.47	.59	.74

Craik and Tulving found that recall was better for positive (yes) responses compared to negative (no) responses, particularly at deeper levels of encoding. They account for this difference by arguing that positive responses provide more elaborate encoding. Thus, performance on recall tests should be higher for positive responses because the word is processed to a greater depth allowing for more elaborate encoding. For example, if the question was "Is the word a type of food?," for the word bear, recall would not have been as good had the word been apple. It is crucial to recognize that we have used an example that requires semantic analysis. Craik and Tulving argue that the difference between yes and no responses shows that analysis of the word's meaning is a necessary aspect of retention. They found significant differences in recall of words encoded with category questions for yes and no responses - 86% and 64% words recalled, respectively. Craik and Tulving also found significant differences in recall between positive and negative responses for rhyme questions - 62% and 42%, respectively. The differences, however, are minimal when positive and negative responses are compared for case questions (less than 1%). Craik and Tulving explain these results by arguing that structural analysis does not require deep encoding for either positive or negative responses. We found similar differences for yes and no responses in our control group, as can be seen in Table 2.

The similar results for response types and yes/no responses for our control group suggest that we have successfully replicated the Craik and Tulving experiment. This successful replication allows us to make valid claims regarding induced-value theory in social science experiments by examining the results for the experimental group.

The Experimental Group
The introduction of induced-value in the experimental group produced significant differences when compared to the results for the control group. These results allow us to reject our null hypothesis: the operations carried out on the material determine retention, not the intention to learn. Thus, a comparison between the control and experimental group of the results shows subjects in social science experiments can be motivated to learn.

The subjects in the experimental group performed better than the control group when the results of all three categories are combined - 73% for the former and 66% for the latter (which is statistically significant at $a = 0.05$). The disparity between the experimental and control group demonstrates that monetary compensation can motivate subjects to learn in experimental trials. While we feel these results are sufficient to prove the validity of our hypothesis about induced-value in social science experiments, further examination of our results will provide greater support for our claim.

Examination of the results for the three response types shows that monetary compensation does affect experimental subjects (Table 3). The subjects in the experimental group performed better on the recall test for all three response types. The difference between the two groups for the rhyme and category results is small, but significant (about 5% greater for the experimental group). The case results, however, are much more significant - about a 12% difference between the experimental and control groups. Craik and Tulving argue that case-encoded words are the hardest to recall. Thus, the difference between the control and experimental groups for case-encoded words suggests that subjects were motivated with money to learn the words presented to them. Another way to analyze this difference is to compare the results for the rhyme-encoded words in the control group and the case-encoded words in the experimental group. The experimental subjects recalled almost as many case-encoded words (about 60%) as the control subjects recalled rhyme-encoded words (about 65%). This suggests that money has become a substitute for the depth of processing for the experimental subjects. In other words, money has motivated subjects to process their observations more deeply.

Table 3
Control vs. Experimental Group
(Proportion of words recalled correctly)

Response Type	Case	Rhyme	Category

Control Group	.48	.65	.84
Experimental Group	.60	.70	.89

The results for yes/no responses in the experimental group are also significantly different when compared to the results for our control group (Table 4). For the experimental group, we found significant differences for all categories regarding yes/no responses. The differences in recall for yes-no responses, however, were relatively smaller for category and rhyme-encoded questions, but larger for case-encoded questions, when compared to the control group. This suggests that monetary compensation motivated subjects to use different strategies.

Table 4
Yes/No Responses for Control and Experimental Groups
(Proportion of Words Recalled Correctly)

Response Type	**Case**	**Rhyme**	**Category**
Our Control Group "YES"	.49	.72	.94
Our Control Group "NO"	.47	.59	.74
Our Experimental Group "YES"	.65	.65	.96
Our Experimental Group "NO"	.55	.65	.83

Conclusion
In order to reject our null hypothesis, we must show that our experiment has achieved the three conditions of the induced-value theory. Had we not successfully achieved all three conditions, we simply would have replicated Craik and Tulving's experiment. As noted earlier we met the first two conditions, monotonicity and salience, in our experimental design.. We proved the difference in results between our control and experimental groups was solely attributable to the two types of payment schemes in our experiment. Thus, our results show that the third condition, dominance, was achieved.

Since dominance was achieved in our experiment, we can now draw two important theoretical implications from the results of our experiment. First, the results for our control and experimental group demonstrate that analysis of the meaning of the words is an important aspect of good retention. Thus, we agree with Craik and Tulving that the depth of processing is a useful framework for understanding retention and recall. Our introduction of induced-value into this classical psychology experiment, however, has led to results that challenge Craik and Tulving's claim that learning is simply a by-product of cognitive processes. Our results suggest that the intention to learn does matter. We do not want to suggest that Craik and Tulving's claim that the operations performed on the material determines retention is wrong. In fact, our results show that the operations performed on the words do matter. We do believe, however, that Craik and Tulving overemphasized the importance of encoding operations. Contrary to the claims of Craik and Tulving, our results suggest that subjects in experimental trials can be motivated to learn. The introduction of a performance-based payment scheme in our experimental group allows us to reject our null hypothesis that the intention to learn does not affect memory recall.

| Cognitive Performance and Economic Incentives Page | The Democratic Dilemma Page | The Related Projects Page |