1995

The effects of feedback and strategy on self-efficacy and computer task performance

Richard John Fogg

Follow this and additional works at: https://scholarworks.lib.csusb.edu/etd-project

Part of the Educational Psychology Commons

Recommended Citation
https://scholarworks.lib.csusb.edu/etd-project/1023

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.
THE EFFECTS OF FEEDBACK AND STRATEGY ON
SELF-EFFICACY AND COMPUTER TASK PERFORMANCE

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Psychology

by
Richard John Fogg
December 1995
THE EFFECTS OF FEEDBACK AND STRATEGY ON
SELF-EFFICACY AND COMPUTER TASK PERFORMANCE

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by
Richard John Fogg
December 1995
Approved by:

Matt L. Riggs, Psychology

Kenneth S. Shultz

Richard S. Trafton

12/5/95
Abstract

The effects of feedback and strategy on self-efficacy and computer task performance were studied utilizing a mixed design. The first independent variable was feedback and the second independent variable was strategy for increasing task performance. Dependent measures were scores on reasoning questions and levels of self-efficacy. The experiment was conducted on personal computers and information elicited from subjects included demographic data, measures of self-efficacy, and performance scores on three series of reasoning questions. Subjects consisted of 104 undergraduate and graduate students from CSUSB.

A computer program was designed to provide bogus performance feedback (i.e., not based on subjects' actual performance) and strategy for improving performance on the reasoning questions. In spite of the bogus nature of the feedback, it included both quantitative (i.e., based on the number of correct responses) and qualitative (i.e., individual subject's performance was compared to other subjects' performance) content in all conditions. After the first two sets of reasoning questions, subjects in all groups received feedback indicating that their performance was average. After completing the last set of reasoning questions, two of the groups received feedback indicating excellent performance; while the remaining two groups received no feedback. In addition, after the second set of
reasoning questions, two groups received strategy designed to improve performance on the last reasoning trial while two groups did not receive the strategy.

It was predicted that: 1) subjects receiving positive feedback after the third reasoning trial would have higher levels of self-efficacy than the groups not receiving the feedback, 2) performance would be positively correlated with subsequent levels of self-efficacy in all conditions; 3) between groups, subjects receiving strategy for improvement on the reasoning trials would increase performance on subsequent tasks as compared to subjects not receiving strategy; 4) between groups, subjects receiving strategy for improving performance would show a positive change in level of self-efficacy; 5) within groups, subjects would show a significant increase in task performance only after they had received strategy for improving their performance; and 6) within groups, after receiving strategy and completing the final task, subjects would show an increase in self-efficacy regardless of whether they received positive feedback; however, the group receiving positive feedback would demonstrate higher levels of self-efficacy than the group not receiving positive feedback.

In spite of one significant result for Hypothesis 6, the data generally failed to support the predictions. Implications of the current study and methodology for future research are discussed.
Acknowledgements

I would like to express my most sincere thanks to my advisor and committee chair, Dr. Matt Riggs, for his continued support and encouragement throughout my graduate studies. My thanks also go to Dr. Kenneth Shultz and Dr. Richard Trafton for their advice and support throughout the course of this project. In a very special way, I wish to express my most heartfelt thanks to Dr. Yu-Chin Chien, Dr. Joanna Worthley, Dr. Kenneth Schultz, and Dr. Matt Riggs for getting me through the "troubles." I owe a special debt of gratitude to Dr. Faith McClure for her unwavering support, listening, and encouragement. Thank you all, my mentors and friends. May peace be with you all.
APPENDIX C
Overview of Self-Efficacy Scales
The Personal Efficacy Scale
The Personal Outcome Expectancy Scale
The Collective Efficacy Scale
The Collective Outcome Expectancy Scale
The Modified Personal Efficacy Scale

APPENDIX D
Reasoning Questions with P-Values

APPENDIX E
Feedback Provided to Subjects

APPENDIX F
Strategy for Improving Performance

APPENDIX G
Design of the Experiment

APPENDIX H
Debriefing Form

APPENDIX I
Questions regarding Subjects Feelings About Study

REFERENCES
List of Figures

Figure 1
Feedback, performance, self-efficacy loop.................11

Figure 2
Mean self-efficacy scores for the third and fourth measures
for groups receiving feedback and groups not receiving
feedback.......................................................35

Figure 3
Mean reasoning scores across trials for groups receiving
strategy and groups not receiving strategy..............39

Figure 4
Mean self-efficacy scores across measures for groups
receiving strategy and groups not receiving strategy.....42

Figure 5
Mean self-efficacy scores across measures for all
subjects.......................................................47

Figure 6
Mean self-efficacy scores across measures for groups
receiving feedback and groups not receiving feedback....50
List of Tables

Table 1
Means, Standard Deviations, Ranges, and Items for the Reasoning Question Trials ............................................. 22

Table 2
Manipulation of the independent variables ...................... 24

Table 3
Means, standard deviations, and alpha levels for the four measures of self-efficacy for all subjects .................. 30

Table 4
Means and standard deviations for performance on reasoning trials for all subjects ......................................... 31

Table 5
Mean self-ratings for performance on reasoning questions for all subjects, and dichotomized by receiving and not receiving strategy and feedback ................................................. 32

Table 6
Mean responses to credibility questions for all subjects, dichotomized by receiving and not receiving strategy and feedback ................................................................. 33

Table 7
Means and standard deviations of levels of self-efficacy for the third and fourth measures for groups receiving and not receiving feedback .................................................. 34
Table 8
ANOVA summary table for results of Hypothesis 1...........36
Table 9
Correlations between reasoning trial scores and levels of
self-efficacy.................................................................37
Table 10
Means and standard deviations for performance on reasoning
trials for groups receiving and not receiving strategy....38
Table 11
ANOVA summary table for results of within-subjects effects
for Hypothesis 3...............................................................40
Table 12
Means and standard deviations for levels of self-efficacy
across measures for groups receiving and not receiving
strategy.................................................................41
Table 13
ANOVA summary table for results of orthogonal contrasts
assessing the effect of strategy for Hypothesis 5...........44
Table 14
ANOVA summary table for results of orthogonal contrasts
without assessing the effect of strategy for Hypothesis
5.................................................................45
Table 15
Means and standard deviations for levels of self-efficacy
across measures for all subjects.................................46
Introduction

A logical goal for organizations is to enhance employee productivity in a cost-effective manner. One traditional mechanism recognized as partially fulfilling that goal is feedback provided via performance evaluations, usually on an annual basis. However, the use of the term "feedback," in and of itself, prohibits a clear understanding of what specifically was done to produce certain effects (Brown, Willis, & Reid, 1981). In addition, one must question the effectiveness of providing feedback on an annual basis. Behaviorists would suggest that such a reinforcement schedule would be of little practical use in altering established behavior patterns (Reynolds, 1968).

Another method of increasing productivity is through the enhancement of employee self-efficacy (i.e., the employee's belief that he or she can perform a certain task). In social cognitive theory, self-efficacy is conceptualized as arising from diverse sources of information conveyed by direct experiences (Bandura, 1982). Performance feedback provides information about prior performance and serves as a basis for evaluating one's capability to perform successfully on subsequent tasks (Bandura, 1986). Many studies have reported significant correlations between self-efficacy and subsequent task performance (e.g., Bandura, 1982; Bandura & Adams, 1977; Bandura, Adams, & Beyer, 1977, Bandura, Adams, Hardy, &
Howells, 1980; Chambliss & Murray, 1979). In studies where efficacy perceptions had been altered by various treatments, the resulting efficacy perceptions still predicted subsequent performance (Gist, 1987). One way to affect levels of self-efficacy is by providing strategy for improving task performance (Locke, Frederick, Lee, & Bobko, 1984).

The purpose of the current study is to investigate the effects of both feedback and strategy on subjects' levels of self-efficacy and performance on a computer task. It is predicted that both treatments (i.e., providing task-specific strategy coupled with timely feedback), will have a direct effect on enhancing levels of self-efficacy and increasing task performance.

**Basic Elements of Feedback and Strategy**

Prue and Fairbank (1981) defined performance feedback as information provided to individuals about the quality and/or quantity of their past performance. Several studies (Ilgen, Fisher, & Taylor, 1979; Landy & Farh, 1980; Larson, 1984; Prue & Fairbank, 1981) have identified various components of this process. According to Duncan and Bruwelheide (1986), one broad viewpoint characterizes feedback as affecting cognitive events. A second framework describes feedback in terms of various behavioral processes.

To fully understand the effects of feedback on both self-efficacy and behavior, one must examine various
components of the "feedback message." Ilgen et al. (1979) identified three dimensions of the "message" that affect perception of feedback: timing, sign, and frequency. "Timing" refers to the interval between the individual's behavior and the receipt of information about that behavior. Ilgen et al. (1979) supported Ammons' (1956) suggestion that the longer the delay in the receipt of feedback, the less effect the feedback would have on subsequent behavior.

Ilgen et al.'s (1979) second dimension was "sign" of feedback, which refers to the positive or negative content of the message. It has been shown (e.g., Feather, 1968; Ilgen, 1971; Ilgen & Hamstra, 1972; Shrauger & Rosenberg, 1970) that positive feedback is perceived and recalled more accurately as a result of its self-image enhancing properties. Negative feedback is neither perceived or recalled accurately because of its failure to bolster self-image. Positive feedback also serves as an incentive, enhances the recipient's perception of competence, and enhances motivated behavior. Negative feedback is said to have the opposite effect (Pavett, 1983).

The third dimension discussed by Ilgen et al. (1979) was "frequency" of feedback. Generally speaking, the recipient is more responsive to frequent feedback than infrequent feedback (Cook, 1968; Ivancevich, Donnelly, & Lyon, 1970). While this could be a function of heightened perception of accuracy, Ilgen et al. (1979) cautioned
against "blindly advocating an increase in feedback frequency, particularly in cases where individuals must interpret complex feedback" (p. 358). This is reasonable because frequent receipt of complex feedback by employees would lead to much time being spent on interpretation of that feedback.

In addition to the feedback message, Ilgen et al. (1979) discussed "source" as a component of the feedback process. They identified a feedback "source" as being: (a) another individual, (b) the task environment, or (c) the individual himself or herself. With regard to feedback message, sources have both credibility and power (Prue & Fairbank, 1981). Credibility is seen as being a function of two factors: first, the source should have the expertise to judge behavior accurately, and second, the source should be perceived as trustworthy. Power is viewed as the ability or official capacity to exercise control or authority. As regards power, the higher the power of the source, the more likely it is that the recipient will alter his or her behaviors in response to the feedback.

It is recognized that the results of the vast majority of studies in the area of feedback, including those discussed thus far, were based on traditional paper-and-pencil measures. However, a few studies have examined changes in task performance as a result of feedback provided via computer.
Earley (1988) studied "computer generated" feedback in a business that consisted of magazine subscriptions sales via telemarketing. A computer program was designed to evaluate the quality (number of errors per order when typed into the computer) and quantity (number of orders entered) of a salesperson's work. His results, in part, showed that subjects had greater belief in the accuracy of computer-generated feedback than in supervisor-generated feedback. He also found that individuals receiving computer-generated feedback had higher self-efficacy expectations than those receiving their feedback from a supervisor.

In a similar study, Northcraft and Earley (1989) simulated stockmarket transactions on a computer and utilized four feedback conditions: organization, supervisor, computer-generated, and self-generated. Organization feedback was "impersonal" feedback in that it was simply information which was slipped under the door of the room where the subject was working. Supervisor feedback was information written by and delivered to the subject by an experimenter confederate. Computer-generated feedback was information based on the success or failure of the stockmarket transactions and could be accessed at will by the subject. Finally, self-generated feedback was also transaction success or failure information but it was maintained on a spreadsheet by the subject. It was found that computer- and self-generated feedback conditions...
produced significantly greater accuracy and timeliness in subsequent performance than did the organization and supervisor generated conditions.

It is evident that timing, sign, frequency, and source of feedback all have an impact on the subsequent behavior of the feedback recipient. This has been shown with both traditional paper-and-pencil measures as well as in computer studies.

**Basic Elements of Self-Efficacy**

Self-efficacy arises from the gradual acquisition of complex cognitive, social, linguistic, and/or physical skills based on experience (Bandura, 1982). Within the work setting, self-efficacy can be viewed as the employee's belief that he or she can perform a certain task. Feedback provides information about prior performance and serves as a basis for evaluating one's capability to perform successfully on subsequent tasks (Bandura, 1986). Bandura and Adams (1977) emphasized that behavior must be measured precisely in the analysis of efficacy and that measures should be tailored to the domain being studied. Bandura (1982) identified four information cues that are thought to influence self-efficacy. From most to least influential, they are (a) enactive mastery, (b) vicarious experience, (c) verbal persuasion, and (d) emotional (physiological) arousal. These cues provide important data, but according to Bandura (1982) it is the cognitive appraisal and
integration of these data that ultimately determine self-efficacy. In the current study, it is the first and the third cues to which attention will be given.

Enactive mastery is defined as repeated performance accomplishments (Bandura, 1982). Mastery is facilitated when gradual accomplishments build the skills, coping abilities, and exposure needed for task performance. Further, while positive mastery experiences increase self-efficacy, negative ones (failures) tend to decrease self-efficacy (Gist, 1987).

The third source of efficacy information is verbal persuasion, which is aimed at convincing a person of his or her capability of performing a task. Verbal persuasion is believed to influence efficacy perceptions in some situations, but it is viewed as less effective than enactive mastery (Bandura, 1982). This is because the experience of success has a more profound impact on an individual than he or she simply being told that they have the ability to perform some task. While verbal persuasion may have the effect of making someone attempt a task, it is the successful completion of that task that truly convinces them of their ability.

As regards the effects of timing on self-efficacy, Bandura (1978) suggested that

to best elucidate the regulatory function of self-referent thought, efficacy judgments and actions
must be measured closely in time. When self-efficacy and actions are measured at widely disparate times, significant intervening experiences might well alter the level of strength of self-judged competence (p. 244).

Thus, relatively frequent feedback delivered shortly after the performance of a task should have the most potent effect on self-efficacy. In the current study, "frequent feedback" is feedback which will be provided immediately after each of the performance trials.

When considering "sign" of feedback in performance on a physical task, Bandura and Cervone (1986) found that unfavorable feedback tended to yield negative self-evaluations and initially increased motivation during subsequent performance on the task. When further feedback indicated that performance continued to fall short of the standard, various reactions were observed across subjects. Some individuals became less motivated; others became demoralized, showing decreased self-efficacy and the selecting of lower goals; while still others remained motivated.

In her review of the organizational implications for self-efficacy, Gist (1987) suggested that there is possibly a reciprocal relationship between feedback and self-efficacy. That is, the various components of the feedback message affect self-efficacy, and levels of self-efficacy
combined with goals affect how the individual responds to feedback. The exact nature of the relationship, however, is open to investigation.

When looking at the effects of sign and source of feedback on self-efficacy and task performance, Fogg (1992) added further support to the contention that computers can serve as credible sources of feedback. In his study, Fogg examined the effects of three different signs of feedback (positive, neutral, negative) and three different sources of feedback (self-generated and delivered via computer, supervisor generated and delivered via computer, supervisor generated and delivered in the traditional written format) on self-efficacy and performance in a cognitive task done on a computer. Although subjects had no measurable differences in baseline levels of self-efficacy, it was shown that changes in levels of self-efficacy could be manipulated by providing bogus feedback to the subjects. Subjects receiving positive feedback demonstrated significant increases in levels of self-efficacy, while negative feedback had the opposite effect. As regards the source of feedback, Fogg found that feedback delivered via the computer was found to be as credible as that delivered in the traditional written format.

**Summary of Feedback and Self-Efficacy**

Taken together, the studies on feedback and self-efficacy suggest that successful individuals receiving
positive feedback on a regular basis shortly after task completion should have the highest levels of self-efficacy and display consistently superior performance behaviors. Lengthening the timing, or altering the sign, or frequency of feedback should result in lower levels of self-efficacy and/or performance. Lower levels of self-efficacy should result in poorer performance and, subsequently, increasingly negative feedback. When thought of in this manner, it is possible to envision a feedback/self-efficacy/performance loop (See Figure 1). (Note: since the current study deals only with an increasing sign of feedback, only the positive end of the spectrum is addressed in the figure.) An increase at any point in the loop would initiate a commensurate increase at the next step in the process. Both feedback "content" and task strategies could serve this enhancement function.
Figure 1

Feedback, performance, self-efficacy loop
Utilization of a credible "source" of feedback in a computer task (e.g., Earley, 1988; Ilgen et al. 1979; Northcraft & Earley, 1989; Prue & Fairbank, 1981), coupled with accurate feedback being provided immediately upon task completion (e.g., Ilgen et al. 1979) should result in greater belief in the accuracy of the feedback and increased motivation to perform in subsequent tasks. Fogg (1992) showed that feedback delivered via computer was perceived to be as credible as that delivered in the traditional written format by a supervisor. Feedback indicating success on the task should be viewed as a mastery experience (Bandura, 1982) and lead to increased levels of self-efficacy. Providing strategy (Locke et al., 1984) as a form of verbal persuasion (Bandura, 1982) to improve performance should result in successful task completion, which in turn should lead to positive feedback and a commensurate increase in self-efficacy. These effects should be measurable using a cognitive task (e.g., Bouffard-Bouchard, 1989), even when done on a computer (e.g., Earley, 1988; Fogg, 1992; Northcraft & Earley, 1989).

Feedback Efficacy Effects

Based on the results of a study on tests of physical exertion, Bandura and Cervone (1986) suggested that motivation based on standards involves a cognitive comparison process. When people commit themselves to explicit standards or goals, the perceived negative
discrepancies between performance and the standard they seek to attain create self-dissatisfaction that serves as a motivational inducement for enhanced effort. Activation of self-evaluative reactions by internal comparison requires both personal standards and knowledge about one's performance level. When these factors are systematically varied, neither knowledge of performance without standards nor standards without knowledge of performance has lasting motivational impact (Bandura & Cervone, 1986). In the absence of absolute standards, individuals must compare themselves with others to determine their ability levels. Individuals are motivated to know their own ability levels because that knowledge enables them to predict their success and avoid failure situations in the future (Farh & Dobbins, 1989).

When faced with a novel task and lacking a clear external criterion by which to evaluate performance, Bandura (1986) stated that "the person is inclined to trust evaluations of her capabilities by those who have access to some objective predictors of performance attainment" (p. 406). Specifically, individuals look for a feedback message "source" that is considered to be credible. That is, individuals seek out information (feedback) about their performance and about the task in order to improve. Positive feedback leads to increases in levels of self efficacy and negative feedback has the opposite effect.
(Fogg, 1992; Pavett, 1983). Thus, in the current study, it is predicted that subjects receiving positive feedback will result in higher levels of self-efficacy.

**Efficacy - Performance**

Bandura (1982) has found perceptions of self-efficacy to be strongly related to both past and future performance, reporting that judgments of efficacy determine "how much effort people will expend and how long they will persist in the face of obstacles of aversive experiences" (p. 123). Whereas negative feedback provides information that an individual's performance has fallen below an acceptable standard, positive feedback indicates that an individual's performance has exceeded an acceptable standard (Bandura, 1986; Ilgen et al. 1979; Latham & Locke, 1991). Thus, positive feedback is likely to enhance self-efficacy, and negative feedback is likely to diminish self-efficacy (Martocchio & Webster, 1992). In addition, research has shown that self-efficacy perceptions are positively related to enhanced problem solving skills in organizational settings (Gist, 1989), thus, it is reasonable to expect that increased levels of self-efficacy will lead to improved performance and, subsequently, more positive feedback.

In her study on the influence of self-efficacy on performance in a cognitive task, Bouffard-Bouchard (1989) found that a high self-efficacy group completed a significantly greater number of problems than did a low
self-efficacy group. In addition, although the students did not differ with respect to the level of problem-solving skills, those who had received positive feedback judged themselves to be more efficacious than those who made their self-appraisals following negative feedback. Bouffard-Bouchard's (1989) findings support Bandura's (1982) contention that self-efficacy expectations contain a motivational component. Based on these findings, it is predicted that performance will be positively correlated with subsequent levels of self-efficacy in all conditions.

**Strategy - Performance**

In a laboratory study, Fogg (1992) was unable to differentiate the effect of strategy from that of practice on task performance. Specifically, subjects in all conditions were given strategy for improving performance on the final task and all subjects demonstrated significant improvement. However, because of an experimental design flaw, Fogg (1992) could not conclude that the improvement was due to receiving strategy, the result of a practice effect, or a combination of both. The design also failed to provide feedback or obtain a measure of self-efficacy after completion of the final task, thus, no conclusions could be drawn about the effect of a mastery experience on subjects' levels of self-efficacy. However, the results of Locke et al. (1984) would suggest that the effect was indeed due to the use of strategy. The current study was a partial
replication of Fogg (1992) and specifically designed to examine the effects of feedback and strategy on self-efficacy and computer task performance. Thus, it was predicted that, between groups, subjects receiving strategy for improvement would increase performance on the final task as compared to subjects not receiving strategy.

**Strategy - Efficacy**

In line with Bandura's (1982) discussion of information cues, one could argue that the offering of task specific strategy is a type of verbal persuasion. Although the individual is not being told he or she has the ability to perform a task per se, they are being provided with the necessary tools to successfully perform a task. These successes are predicted to increase motivation on future tasks and lead to additional mastery experiences, thus, increasing efficacy beliefs and performance. Thus, in the current study, it was predicted that, between groups, subjects receiving strategy for improving performance would show a positive change in level of self-efficacy.

**Efficacy vs. Strategy - Performance**

In a laboratory setting, Locke et al. (1984) investigated the effects of goals, self-efficacy, and task strategies on goal choice and task performance over repeated trials. A path analysis indicated that performance was affected by, among other things, both strength of self-efficacy and strategies used. Locke et al.'s (1984) study
suggests that providing specific feedback and task specific strategy serves to enhance performance. Subsequent evaluation of that performance should result in higher scores and higher levels of self-efficacy.

While analysis of feedback in terms of behavioral outcomes is important for the practitioner, it is not reasonable to assume a priori that feedback is a reinforcer (Duncan & Bruwelheide, 1986). Locke's (1968) goal-setting theory states that feedback only affects behavior when it leads the recipient to set a goal. In this sense, feedback provides information so the recipient can set a goal for himself or herself, or recognize a discrepancy between his or her behavior and an existing standard. Thus, improvements in performance can be examined in terms of whether the individual sets new goals based on feedback received. In a similar manner, the enhancement of self-efficacy is also incremental (Bandura, 1982). Based on this information, it was predicted that, within groups, subjects would show a significant increase in task performance only after they received strategy for improving their performance. It was also predicted that after receiving strategy and completing the final task, subjects would show an increase in self-efficacy regardless of whether they received feedback; however, the group receiving positive feedback would demonstrate higher levels of self-efficacy than the group not receiving positive feedback.
Summary of Hypotheses

Based on the review of the literature, it was predicted that:

1) Subjects receiving positive feedback after the third reasoning trial would have higher levels of self-efficacy than subjects not receiving the feedback.

2) Performance would be positively correlated with subsequent levels of self-efficacy in all conditions.

3) Between groups, subjects receiving strategy for improvement on the reasoning trials would increase performance on the final task as compared to subjects not receiving strategy.

4) Between groups, subjects receiving strategy for improving performance would show a positive change in level of self-efficacy.

5) Within groups, subjects would show a significant increase in task performance only after they had received strategy for improving their performance.

6) Within groups, after completing the final task, subjects would show an increase in self-efficacy regardless of whether they received positive feedback; however, the group receiving positive feedback would demonstrate higher levels of self-efficacy than the group not receiving positive feedback.
Method

Subjects

One hundred and four students (39 male and 65 female) from a medium sized Southwestern state university participated as subjects in the experiment. The average age of the subjects was 27.57 years. Subjects were recruited from psychology classes and, upon request, were given extra credit slips for their participation. All subjects were told they were participating in a computer-generated testing program which was being evaluated for possible use on the campus. All subjects were treated in accordance with Principle 9 of the "Ethical Principles of Psychologists" (American Psychological Association, 1982).

Apparatus

The experiment was conducted in the Psychology Department computer laboratory. Eight personal computers were utilized for the study so that eight subjects could be run simultaneously. The experimental program was administered using MicroCAT Testing System software, a microcomputer-based system for developing, administering, scoring, and analyzing computerized tests (Assessment Systems Corporation, 1989). With this program it was possible to write all items, instruct the subject as to what would be required during his or her participation, present questions to the subject, have the subject either enter a numerical response or select an available multiple choice
response, and record all of the subject's responses on a floppy disk.

All subjects were asked to read and sign an Informed Consent Form (See Appendix A). The questions asked of the subjects fell into three general categories: demographic, self-efficacy, and reasoning. Demographic data requested from subjects included the last four digits of their student number (case ID for computer analysis), age, ethnicity, and sex (See Appendix B for demographic questions). To reinforce the mundane realism of the study, subjects were asked if they had ever taken computer-generated tests of any kind, plus five other questions of that nature (See Appendix B). This was followed by the first measure of self-efficacy.

The self-efficacy measure used was the personal efficacy component of the Riggs and Knight (in press) Efficacy Scale (See Appendix C for efficacy scale). This scale has 10 questions and uses a six point Likert-type response format ranging from 1 "Strongly Agree" to 6 "Strongly Disagree". Each subject's level of self-efficacy was measured four times during the course of the study. The initial measure was used as a baseline.

The task itself consisted of responding to three series of general reasoning problems (See Appendix D for questions and item difficulty levels). The questions were standard cognitive reasoning questions adapted from Sternberg (1986).
The reasoning questions were pilot-tested by Fogg (1992) using a paper-and-pencil format. That was followed by an item analysis to assess item difficulty (p-values) for each question. In making up the sample questions and the three series of questions for the tasks in the current study, sets of questions were matched for p-values (See Table 1).
Table 1.
Means, Standard Deviations, Ranges, and Items for the Reasoning Question Trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>$M_p$</th>
<th>SD</th>
<th>Range</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Items</td>
<td>.84</td>
<td>.05</td>
<td>.80-.91</td>
<td>30, 33, 2, 26</td>
</tr>
<tr>
<td>1st</td>
<td>.42</td>
<td>.16</td>
<td>.18-.78</td>
<td>15, 6, 36, 9, 37, 22, 27, 40, 32, 11, 39, 35</td>
</tr>
<tr>
<td>2nd</td>
<td>.42</td>
<td>.18</td>
<td>.11-.77</td>
<td>16, 19, 14, 38, 10, 12, 7, 20, 4, 8, 29, 24</td>
</tr>
<tr>
<td>3rd</td>
<td>.41</td>
<td>.17</td>
<td>.12-.78</td>
<td>3, 23, 1, 28, 5, 13, 34, 21, 17, 25, 31, 18</td>
</tr>
</tbody>
</table>

Item numbers are listed in the order in which they were presented to subjects. The item numbers correspond to the numbers in Appendix D.

Based on the results of the pilot testing (Fogg, 1992), the four easiest questions were used for the sample questions. It was hoped that by using the four easiest questions understanding for the subjects would be enhanced.
and their chances of doing well on the scored reasoning trials would increase. The remaining 36 questions were divided into sets of 12 for each of the three reasoning trials. Taken together, the 36 questions had a mean p-value of .42 (SD = .17) with a range of .67 (.78 to .11). All reasoning question responses were multiple choice with four response options.

After each set of reasoning questions a screen appeared which asked subjects to type in a number which represented the percentage of questions they felt they had answered correctly.

Two different types of feedback were used during the course of the study, neutral and positive. The first two administrations of feedback were neutral (See Appendix E for exact feedback wording). After the third set of reasoning questions, two groups of subjects received positive feedback. The content of the feedback messages was identical in wording for all subjects. Subjects were also told their performance was being compared to that of other subjects who participated in the study.

Subjects in two groups were given strategy that provided accurate information as to how they could improve their scores on subsequent reasoning trials (See Appendix F). The strategy was modeled on that suggested by Sternberg (1986).
Design

The study utilized a 2 (feedback) x 2 (strategy) x (3) (reasoning trials) x (4) (self-efficacy measures) mixed design. The first independent variable was feedback and the second independent variable was strategy for increasing task performance. Both the first and the second independent variables were discrete. Table 2 shows how the feedback and strategy were presented to the four groups of subjects.

Table 2.
Manipulation of the independent variables

<table>
<thead>
<tr>
<th>No Strategy</th>
<th>Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Group 2</td>
</tr>
<tr>
<td>Group 3</td>
<td>Group 4</td>
</tr>
</tbody>
</table>
The two repeated measures used in this study were the scores on the three reasoning trials, and the scores on the four administrations of the self-efficacy scale. The discrete variables were combined with the repeated variables as follows: 2 (feedback) x (4) (self-efficacy scores); 2 (strategy) x (4) (self-efficacy scores); and 2 (strategy) x (3) (reasoning trial scores).

For each subject, the experiment took place in one session. Subjects were randomly assigned to one of four experimental groups (See Appendix G) with feedback and strategy being manipulated: Group 1 received both the strategy component and the positive feedback after the third reasoning trial, Group 2 did not receive the strategy but received the positive feedback, Group 3 received the strategy but not the positive feedback after the third reasoning trial, and Group 4 received neither the strategy nor the positive feedback. Number of subjects per group was as follows: Group 1 = 25, Group 2 = 26, Group 3 = 28, and Group 4 = 25.

**Procedure**

The experiment took place in the computer laboratory of the Department of Psychology over a one week period, and subjects were recruited from both undergraduate and graduate psychology courses. Appointments were not required; subjects participated when they arrived at the laboratory. After signing an informed consent, subjects were told
to sit at a computer station. All directions for participating in the study were delivered via IBM compatible personal computers. With the exception of five answers requiring numbers, all items utilized a format in which the response was highlighted by pressing the "spacebar" and recorded by pressing the "enter" key. Numerical responses required the subject to type in a number and then press the "enter" key to record the response. When the response was recorded, the next screen appeared. It was an untimed study, however, the average time of participation was approximately twenty minutes.

All sessions began with directions on how to enter responses on the computer, and each subject had a chance to practice that skill. The first recorded responses were the demographic questions. They were followed by six questions regarding prior computer testing experience, which in turn were followed by the baseline measure of self-efficacy. Next were the four sample reasoning questions and instructions on how to complete the task. The first trial of reasoning questions followed immediately after the samples. Upon completion of the first set of reasoning questions, a screen appeared telling the subjects to wait while the computer scored their responses. The same screen appeared upon completion of each set of reasoning questions; however, the length of time the message appeared on the screen varied to provide the appearance that the computer
was actually calculating their score. Subjects were then asked to enter a number to represent the percentage of questions they felt they had answered correctly. Neutral feedback was then given to the subjects for the first time. This was followed by the second measure of self-efficacy.

At the end of the second measure of self-efficacy a second series of reasoning questions, the same in nature as the first, began. This was followed by a second request for estimated percentage of correct responses, a second delivery of neutral feedback, and a third measure of self-efficacy. Subjects in Group 1 and 3 were then given strategy which provided accurate information as to how they could improve their scores on subsequent reasoning trials (See Appendix E). A third and final series of reasoning questions was given, a final request for the estimated percentage of correct responses, and this was followed by positive feedback for only Groups 1 and 2. Groups 3 and 4 received no feedback after the final set of reasoning questions. This was followed by a final (fourth) measure of self-efficacy. Four final questions regarding their feelings about different aspects of the study were asked of all subjects (See Appendix I for the exact wording of these questions).

The order of presentation was the same for Group 2 except that no strategy was provided (See Table 2). Group 3 received no positive feedback after the third reasoning trial. Group 4 received neither the strategy nor the
positive feedback.

At the conclusion of the experiment all subjects were debriefed regarding the true nature of the experiment and the amount and type of deception that was utilized (See Appendix H). This included the fact that neutral feedback was given after the first and second reasoning trials and positive feedback was provided after the third reasoning trial for Groups 1 and 2, regardless of their actual performance. It was not anticipated that the bogus feedback would cause any undue distress due to the sign of the feedback (i.e., neutral and positive).
Results

The means and standard deviations for levels of self-efficacy for all subjects can be seen in Table 3. The scale used for the self-efficacy measure is provided at the bottom of the table. Note that lower scores reflect higher levels of self-efficacy. An analysis was conducted to test reliability of the self-efficacy scale for this study. The Cronbach's Alpha reliability coefficients are also provided in Table 3. The levels of Cronbach's Alpha range from .84 to .88, and became progressively higher across administrations.
Table 3
Means, standard deviations, and alpha levels for the four measures of self-efficacy for all subjects

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>2.70b</td>
<td>0.83</td>
<td>.84</td>
</tr>
<tr>
<td>2nd</td>
<td>3.00</td>
<td>0.84</td>
<td>.85</td>
</tr>
<tr>
<td>3rd</td>
<td>2.97</td>
<td>0.84</td>
<td>.86</td>
</tr>
<tr>
<td>4th</td>
<td>2.90</td>
<td>0.87</td>
<td>.88</td>
</tr>
</tbody>
</table>

a Possible score range 1 to 6.
b Scale used was as follows:

1 - Strongly Agree  4 - Disagree Somewhat
2 - Agree         5 - Disagree
3 - Agree Somewhat 6 - Strongly Disagree
The means and standard deviations for the three reasoning trials for all subjects are given in Table 4. It can be seen that the scores for all trials were less than 50% of the possible 12 correct for each trial.

Table 4  
Means and standard deviations for performance on reasoning trials for all subjects

<table>
<thead>
<tr>
<th>Reasoning Trials</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>1st</td>
<td>4.94 (2.23)</td>
<td>4.85 (2.03)</td>
<td>5.34 (2.29)</td>
</tr>
</tbody>
</table>

*a Possible score range 0 to 12
The results of the self-ratings for performance can be seen in Table 5. In all cases the average self-ratings were lower than the scores provided in the feedback given to subjects (i.e., 68% correct for Trial 1, 70% correct for Trial 2, and, for Groups 1 and 2, 94% correct for Trial 3). However, the self-ratings were also higher in all cases than actual performance as indicated in Table 4.

Table 5

Mean self-ratings for performance on reasoning questions for all subjects, and dichotomized by receiving and not receiving strategy and feedback

<table>
<thead>
<tr>
<th>Reasoning Trials</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Subjects</td>
<td>65.42</td>
<td>64.60</td>
<td>65.44</td>
</tr>
<tr>
<td>Rec'd Strategy</td>
<td>65.00</td>
<td>64.42</td>
<td>64.35</td>
</tr>
<tr>
<td>No Strategy</td>
<td>65.84</td>
<td>64.78</td>
<td>66.60</td>
</tr>
<tr>
<td>Rec'd Pos. Feedback</td>
<td>64.92</td>
<td>63.16</td>
<td>64.94</td>
</tr>
<tr>
<td>No Pos. Feedback</td>
<td>65.90</td>
<td>65.98</td>
<td>65.92</td>
</tr>
</tbody>
</table>

*All figures are given as estimated percent correct.*
Mean responses to the questions regarding the credibility of the study (See Appendix I) are provided in Table 6.

Table 6

Mean responses to credibility questions for all subjects, dichotomized by receiving and not receiving strategy and feedback

<table>
<thead>
<tr>
<th>Credibility Questions</th>
<th>Strategy Helpful</th>
<th>Feedback Credible</th>
<th>Feedback Accurate</th>
<th>Deception Was Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Subjects</td>
<td>2.81</td>
<td>2.94</td>
<td>2.88</td>
<td>3.53</td>
</tr>
<tr>
<td>Strategy</td>
<td>2.51</td>
<td>2.81</td>
<td>2.83</td>
<td>3.60</td>
</tr>
<tr>
<td>No Strategy</td>
<td>3.12</td>
<td>3.08</td>
<td>2.94</td>
<td>3.47</td>
</tr>
<tr>
<td>Pos. Feedback</td>
<td>2.55</td>
<td>2.90</td>
<td>3.02</td>
<td>3.26</td>
</tr>
<tr>
<td>No Pos. Feedback</td>
<td>3.06</td>
<td>2.98</td>
<td>2.75</td>
<td>3.79</td>
</tr>
</tbody>
</table>

* Scale used was as follows:

1 - Strongly Agree
2 - Agree
3 - Agree Somewhat
4 - Disagree Somewhat
5 - Disagree
6 - Strongly Disagree
The means and standard deviations for levels of self-efficacy for the third and fourth measures for those receiving positive feedback after the third reasoning trial (Groups 1 and 2) and those not receiving positive feedback after the third reasoning trial (Groups 3 and 4) can be seen in Table 7. (Note: lower scores = higher self-efficacy.)

<table>
<thead>
<tr>
<th>Feedback</th>
<th>3rd M (SD)</th>
<th>4th M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>2.94 (.76)</td>
<td>2.75 (.76)</td>
</tr>
<tr>
<td>No</td>
<td>3.00 (.92)</td>
<td>3.05 (.94)</td>
</tr>
</tbody>
</table>

* Possible score range 1 to 6.

A graph of the mean self-efficacy scores for the third and fourth measures can be seen in Figure 2.
The first hypothesis of the study predicted that subjects receiving positive feedback after the third reasoning trial would have higher levels of self-efficacy than subjects not receiving the feedback. To test the hypothesis, an ANOVA was conducted on the fourth measure of self-efficacy, with positive feedback and strategy as the grouping variables. The means and standard deviations on the fourth measure of self-efficacy for the subjects
receiving feedback (Yes) and the group not receiving feedback (No) can be seen in Table 7. The summary table for the ANOVA can be seen in Table 8.

Table 8
ANOVA summary table for results of Hypothesis 1

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback</td>
<td>2.35</td>
<td>1</td>
<td>2.35</td>
<td>3.22</td>
<td>.08</td>
</tr>
<tr>
<td>Strategy</td>
<td>0.01</td>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>.91</td>
</tr>
<tr>
<td>FB(^1) x ST(^2)</td>
<td>2.25</td>
<td>1</td>
<td>2.25</td>
<td>3.09</td>
<td>.08</td>
</tr>
<tr>
<td>Residual</td>
<td>72.92</td>
<td>100</td>
<td>0.73</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) FB = Feedback
\(^2\) ST = Strategy

It can be seen from the results of the summary table that there was no main effect for feedback, \(F(1,100) = 2.35, p=.08\), and there was no feedback by strategy interaction \(F(1,100) = 2.25, p=.08\); thus, the first hypothesis was not supported.

The results of the correlations to test the second hypothesis can be seen in Table 9. It was predicted that performance would be positively correlated with subsequent levels of self-efficacy in all conditions. A significant
negative correlation was found between the first reasoning trial and the second measure of self-efficacy, \( r(103) = -0.29, p=0.002 \), showing a strong relationship for these two measures. However, all other correlations were not significant, thus, the second hypothesis was not supported.

Table 9
Correlations between reasoning trial scores and levels of self-efficacy

<table>
<thead>
<tr>
<th>Self-Efficacy Measures</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning Trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 1</td>
<td>-.29*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 2</td>
<td></td>
<td>-.06NS</td>
<td></td>
</tr>
<tr>
<td>Trial 3</td>
<td></td>
<td></td>
<td>-.06NS</td>
</tr>
</tbody>
</table>

* \( p<.01 \)

The means and standard deviations for performance on reasoning trials for groups receiving strategy to improve performance on the third reasoning trial (Groups 1 and 3) and for groups not receiving strategy to improve performance on the third reasoning trial (Groups 2 and 4) are given in Table 10. As can be seen in the table, the groups receiving strategy had a mean score of 5.51 for performance on the
third reasoning trial while the groups which did not receive the strategy had a means score of 5.16.

**Table 10**

**Means and standard deviations for performance on reasoning trials for groups receiving and not receiving strategy**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>1st M (SD)</th>
<th>2nd M (SD)</th>
<th>3rd M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>5.09 (2.39)</td>
<td>4.81 (1.79)</td>
<td>5.51 (2.45)</td>
</tr>
<tr>
<td>No</td>
<td>4.78 (2.08)</td>
<td>4.88 (2.27)</td>
<td>5.16 (2.11)</td>
</tr>
</tbody>
</table>

*Possible score range 0 to 12*

A graph of the mean reasoning scores across trials can be seen in Figure 3.
Figure 3

Mean reasoning scores across trials for groups receiving strategy and groups not receiving strategy

Reasoning Scores by Strategy

Hypothesis 3 predicted that, between groups, subjects receiving strategy for improvement on the reasoning trials would increase performance on the final task as compared to subjects not receiving strategy. A repeated measures ANOVA was conducted to examine the effects of strategy and feedback on reasoning scores from the second to the third reasoning trials. The summary table for the ANOVA can be seen in Table 11.
Table 11
ANOVA summary table for results within-subjects effects for Hypothesis 3

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning</td>
<td>11.90</td>
<td>1</td>
<td>11.90</td>
<td>5.06</td>
<td>.03</td>
</tr>
<tr>
<td>FB x REAS</td>
<td>3.02</td>
<td>1</td>
<td>3.02</td>
<td>1.28</td>
<td>.26</td>
</tr>
<tr>
<td>ST x REAS</td>
<td>2.10</td>
<td>1</td>
<td>2.10</td>
<td>.89</td>
<td>.35</td>
</tr>
<tr>
<td>FB x ST x REAS</td>
<td>.32</td>
<td>1</td>
<td>.32</td>
<td>.14</td>
<td>.71</td>
</tr>
<tr>
<td>Residual</td>
<td>235.29</td>
<td>100</td>
<td>2.35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[^ FB = Feedback, ^ REAS = Reasoning, ^ ST = Strategy

There was a significant change in reasoning scores between Trial 2 and Trial 3, \( F(1,102) = 5.06, p=.03 \); however, there was no effect for strategy, thus, the third hypothesis was not supported.

The means and standard deviations for levels of self-efficacy across measures for groups receiving strategy to improve performance on the third reasoning trial (Groups 1 and 3) and for groups not receiving strategy to improve performance on the third reasoning trial (Groups 2 and 4) can be seen in Table 12.
Table 12

Means and standard deviations for levels of self-efficacy across measures for groups receiving and not receiving strategy

<table>
<thead>
<tr>
<th>Strategy</th>
<th>1st M (SD)</th>
<th>2nd M (SD)</th>
<th>3rd M (SD)</th>
<th>4th M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>2.74 (.87)</td>
<td>3.00 (.88)</td>
<td>2.97 (.83)</td>
<td>2.90 (.83)</td>
</tr>
<tr>
<td>No</td>
<td>2.65 (.79)</td>
<td>3.00 (.80)</td>
<td>2.98 (.86)</td>
<td>2.91 (.91)</td>
</tr>
</tbody>
</table>

* Possible score range 1 to 6

A graph of the mean self-efficacy scores across measures can be seen in Figure 4.
An independent groups t-test was conducted to test the fourth hypothesis which predicted that, between groups, subjects receiving strategy for improving performance would show a positive change in level of self-efficacy. The mean and standard deviation on the fourth measure of self-efficacy for the groups receiving strategy to improve performance was 2.90 (.83) and for the groups not received strategy was 2.91 (.91). The results of the t-test showed no effect for strategy, \( t(102) = .05, p = .964 \); thus, the
fourth hypothesis was not supported.

Hypothesis five predicted that, within groups, subjects would show a significant increase in task performance only after they received strategy for improving their performance (See Table 10 for a summary of means and standard deviations for performance on reasoning trials for groups receiving and not receiving strategy, and Figure 3 for a graph of the means). Repeated measures ANOVAs and specific contrasts across trials (i.e., 1 vs. 2, and 1 & 2 vs. 3), both assessing and not assessing the effect of strategy, were conducted to test the hypothesis. The first contrast was conducted to test whether there was any increase in performance prior to the administration of the strategy. The second contrast was to actually test the prediction that an increase in performance would only occur after subjects received the feedback for improving performance. A summary table of the results of the repeated measures ANOVA which assessed the effects of strategy can be seen in Table 13.
Table 13

ANOVA summary table for results of orthogonal contrasts assessing the effect of strategy for Hypothesis 5

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2</td>
<td>1.89</td>
<td>1</td>
<td>1.89</td>
<td>.77</td>
<td>.38</td>
</tr>
<tr>
<td>Residual</td>
<td>249.63</td>
<td>102</td>
<td>2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 &amp; 2 vs. 3</td>
<td>.94</td>
<td>1</td>
<td>.94</td>
<td>.32</td>
<td>.57</td>
</tr>
<tr>
<td>Residual</td>
<td>299.49</td>
<td>102</td>
<td>2.93</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the orthogonal contrasts assessing the effect of strategy showed no significant within group difference between Trial 1 versus Trial 2 ($F(1, 102) = .77$, $p=.38$) as was expected. However, there was also no significant difference between Trials 1 and 2 versus Trial 3 ($F(1, 102) = .32$, $p=.57$), thus, the fifth hypothesis was not supported.

The summary table of the results of a repeated measures ANOVA examining the same orthogonal contrasts without assessing the effects of strategy can be seen in Table 14.
Table 14

ANOVA summary table for results of orthogonal contrasts without assessing the effect of strategy for testing

Hypothesis 5

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2</td>
<td>.44</td>
<td>1</td>
<td>.44</td>
<td>.18</td>
<td>.67</td>
</tr>
<tr>
<td>Residual</td>
<td>249.63</td>
<td>102</td>
<td>2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 &amp; 2 vs. 3</td>
<td>13.42</td>
<td>1</td>
<td>13.42</td>
<td>4.57</td>
<td>.04</td>
</tr>
<tr>
<td>Residual</td>
<td>299.49</td>
<td>102</td>
<td>2.93</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the orthogonal contrasts without assessing the effect of strategy showed no significant within group difference between Trial 1 versus Trial 2 ($F(1, 102) = .18, p=.67$) as was expected. There was, however, a significant difference between Trials 1 and 2 versus Trial 3 ($F(1, 102) = 4.57, p=.04$). This result indicates that there was a significant increase in reasoning scores for Trial 3; however, because it was not due to the strategy as predicted, the hypothesis was not supported.

The sixth and last hypothesis predicted that, within groups, after completing the final task, subjects would show an increase in the fourth measure of self-efficacy regardless of whether they received positive feedback; however, the group receiving positive feedback would
demonstrate higher levels of self-efficacy than the group not receiving positive feedback.

The means and standard deviations for levels of self-efficacy (across measures) for all subjects can be seen in Table 15 (Note: lower scores reflect higher levels of self-efficacy). It can be seen that subjects had higher levels of self-efficacy (M = 2.70) before treatment than they did for any of the subsequent measures.

Table 15
Means and standard deviations for levels of self-efficacy across measures for all subjects

<table>
<thead>
<tr>
<th>Self-Efficacy Measure</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>2.70 (.83)</td>
<td>3.00 (.84)</td>
<td>2.97 (.84)</td>
<td>2.90 (.87)</td>
<td></td>
</tr>
</tbody>
</table>

a Possible score range 1 to 6.

A graph of the mean self-efficacy scores across measures can be seen in Figure 5.
To test the first part of the hypothesis a repeated measures ANOVA was conducted for all subjects, across administrations and a significant within-subjects effect was found, $F(3, 306) = 13.80, p<.0005$. To identify where the changes occurred orthogonal contrasts of the self-efficacy measures were conducted and the results can be seen in Table 16.
Table 16

ANOVA summary table for results of orthogonal contrasts for all subjects on the four self-efficacy measures

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2</td>
<td>4.89</td>
<td>1</td>
<td>4.89</td>
<td>24.08</td>
<td>.000</td>
</tr>
<tr>
<td>Residual</td>
<td>20.92</td>
<td>102</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 vs. 3</td>
<td>1.09</td>
<td>1</td>
<td>1.09</td>
<td>10.75</td>
<td>.001</td>
</tr>
<tr>
<td>Residual</td>
<td>10.46</td>
<td>102</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 3 vs. 4</td>
<td>.01</td>
<td>1</td>
<td>.01</td>
<td>.07</td>
<td>.784</td>
</tr>
<tr>
<td>Residual</td>
<td>14.22</td>
<td>102</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It can be seen from the summary table in Table 16 and the mean self-efficacy scores in Table 15 that there was a significant drop in self-efficacy from the first to the second administration (M = 2.70 vs. 3.00). Another significant drop in self-efficacy was seen for the third administration (M = 2.85 vs. 2.97). However, there was no change seen with the last administration (M = 2.89 vs. 2.90). The first two changes in level of self-efficacy were not in the predicted direction, and there was no change seen for the final measure of self-efficacy, thus, the first part of Hypothesis 6 was not supported.

The means and standard deviations for levels of self-efficacy (across measures) for those receiving positive
feedback after the third reasoning trial (Groups 1 and 2) and those not receiving positive feedback after the third reasoning trial (Groups 3 and 4) can be seen in Table 17.

Table 17
Means and standard deviations for levels of self-efficacy across measures for groups receiving and not receiving feedback

<table>
<thead>
<tr>
<th>Feedback</th>
<th>1st M (SD)</th>
<th>2nd M (SD)</th>
<th>3rd M (SD)</th>
<th>4th M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>2.69 (.89)</td>
<td>2.99 (.81)</td>
<td>2.94 (.76)</td>
<td>2.75 (.76)</td>
</tr>
<tr>
<td>No</td>
<td>2.71 (.77)</td>
<td>3.01 (.88)</td>
<td>3.00 (.92)</td>
<td>3.05 (.94)</td>
</tr>
</tbody>
</table>

* Possible score range 1 to 6.

A graph of the mean self-efficacy scores across measures can be seen in Figure 6.
To test the second part of the hypothesis specific contrasts again had to be conducted on the self-efficacy measures (i.e., 1 vs. 2, 1 2 vs. 3, and 1 2 3 vs. 4); however, these contrasts took into account the effect of feedback. The results of the orthogonal contrasts can be seen in Table 18.
Table 18

Results of orthogonal contrasts on self-efficacy measures for subjects receiving and not receiving feedback

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2</td>
<td>.002</td>
<td>1</td>
<td>.002</td>
<td>.01</td>
<td>.913</td>
</tr>
<tr>
<td>Residual</td>
<td>20.92</td>
<td>102</td>
<td>.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 vs. 3</td>
<td>.02</td>
<td>1</td>
<td>.02</td>
<td>.20</td>
<td>.653</td>
</tr>
<tr>
<td>Residual</td>
<td>10.44</td>
<td>102</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2 3 vs. 4</td>
<td>1.34</td>
<td>1</td>
<td>1.34</td>
<td>10.63</td>
<td>.002</td>
</tr>
<tr>
<td>Residual</td>
<td>12.88</td>
<td>102</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since the positive feedback was provided after the third reasoning trial, the last contrast was of interest to test the effect of positive feedback on self-efficacy. As can be seen in the Table 17, there was a significant difference between measures 1, 2, and 3 versus 4. Since the second part of the hypothesis predicted that there would be a greater increase in self-efficacy for groups receiving feedback, the sixth hypothesis was supported.
Discussion

The purpose of this study was to examine the relationships between strategy to improve performance and performance feedback on subjects' levels of self-efficacy and their actual performance on a novel computer task. It was postulated that a "performance loop" exists (See Figure 1) in which positive feedback leads to higher levels of self-efficacy which, in turn, results in better performance. Furthermore, an increase at any one point (i.e., feedback, self-efficacy, or performance) would result in a concomitant increase at successive points in the loop. Although the presence of this type of performance loop is well grounded theoretically, the current study generally failed to support its development.

Our first hypothesis predicted that subjects receiving positive feedback after the third reasoning trial would have higher levels of self-efficacy than subjects not receiving the positive feedback; however, our analysis failed to support this prediction. It is possible that this was a function of the perceived credibility of the feedback provided to the subjects. Previous studies (Bandura & Cervone, 1983, 1986) had subjects' perform physical tasks and they then provided accurate feedback to subjects. Because the tasks were physical, the subjects had some intuitive notion of their level of success; and because the feedback provided was roughly in line with their intuitive
notions it was perceived to be credible.

In the current study, it was thought that the use of a novel task (answering cognitive reasoning questions on a computer) would somewhat limit the subjects' abilities to intuitively assess their performance, thus, making the feedback more believable. However, that did not appear to be the case. In examining the responses to the questions regarding the subjects' perceived credibility of the manipulations used in the study (Table 6, p. 33), it was found that the subjects who received the positive feedback after the third reasoning trial rated that feedback only slightly more in the anticipated direction (M = 2.90) than those who received only the neutral feedback (M = 2.98). It is conceivable that being told their performance changed from "about average" to "excellent" was not in keeping with their intuition about their performance, thus, their self-efficacy ratings did not increase.

Support for this premise can be seen in the subjects' self-ratings of their performance (Table 5, p. 32). For example, after the third reasoning trial, subjects in Groups 1 and 2 (those receiving positive feedback) had an average self-rating of 64.94% and subjects in Groups 3 and 4 (those not receiving positive feedback) had an average self-rating of 65.92%. We can see that the self-ratings for Groups 1 and 2 were not in line with the feedback that their performance was at the 94% level.
In keeping with the notion of a "performance loop," the second hypothesis predicted that performance would be positively correlated with subsequent levels of self-efficacy in all conditions. This prediction was not supported. Looking again at the differences between self-ratings of performance and feedback (Table 5, p. 32), one might postulate that the discrepancy is a function of this factor. Across all subjects, the mean self-rating scores for the three reasoning trials were 65.42%, 64.60%, and 65.44%; while actual performance scores were 41.19%, 40.38%, and 44.47%. That represented difference-scores of 24.33%, 24.22%, and 20.97% respectively. Self-ratings were inflated over 20% for all trials and the feedback provided to subjects inflated those figures even more. Thus, the accuracy and credibility of the feedback were compromised and performance suffered as a result (Ilgen et al. 1979), as did self-efficacy (Gist, 1987).

The third hypothesis predicted that, between groups, subjects receiving strategy for improvement on the reasoning trials would increase performance on the final task as compared to subjects not receiving strategy. Once again, the hypothesis was not supported. Although there was no statistical difference between groups, the mean score for the group receiving strategy was more in the anticipated direction than the mean for the group that did not receive the strategy (M = 5.51 v. 5.16).
When examining the results of the question "I feel that the strategy provided for improving my performance was helpful," on a scale from 1 (Strongly Agree) to 6 (Strongly Disagree), the mean rating was 2.51 for those who received the strategy and 3.12 for those who did not receive the strategy. One question that remains unanswered is, what strategy did those who did not receive the "Strategy for Improving Performance" screens think they received? Perhaps the groups that did not receive specific strategy (Groups 1 and 3) believed that the general instructions and examples of reasoning questions given to all subjects before the actual reasoning trials constituted the strategy to which the question referred. This, however, is purely speculation because it was not empirically examined.

Regardless of where the subjects believed the strategy originated, their perceptions of the usefulness of the strategy was certainly in keeping with the fact that there was no difference in reasoning trial scores between the groups receiving and not receiving the strategy. This result, however, was not consistent with Fogg (1992) where all subjects, regardless of the sign of the feedback (positive, neutral, negative), showed an increase in performance after receiving task specific strategy for improvement. Therefore, the difference in results between the two studies remains an open question.

Another expected effect for strategy was reflected in
the fourth hypothesis which predicted that, between groups, subjects receiving strategy for improving performance would show a positive change in level of self-efficacy. The results failed to show the predicted effect. The obtained result also failed to support Locke et al. (1984) who suggested that providing task strategies should not only improve task performance but also strengthen levels of self-efficacy.

Although not empirically examined, it is possible that, because the subjects did not rate the strategy that was provided as particularly helpful, they did not believe that it was particularly instrumental in increasing their performance. Therefore, since it did not result in perceptions of an increase in task performance, it would also not have had the concomitant effect of raising levels of self-efficacy.

Future studies should examine the strategy effect in isolation. Specifically, future studies should provide strategy to one group, and then determination of the effects of the strategy should be based solely on behavioral responses (i.e., responses on subsequent measures of self-efficacy), rather than assessment of the subjects' beliefs that the strategy helped their performance. It is possible that by asking a question about the usefulness of the strategy those who received the strategy believed they did not use it to their best advantage, and those who did not
receive the strategy believed that they missed it completely; thus, self-efficacy went down because the subjects felt like they had either done something wrong or failed completely in their mission.

The fifth hypothesis predicted that, within groups, subjects would show a significant increase in task performance only after they had received strategy for improving their performance. The results failed to support this hypothesis. Although there was no significant difference between groups, the means for the two groups were in the expected directions (i.e., strategy group: M = 5.51, no strategy group: M = 5.16). Turning once again to Table 10 (p. 38), it can be seen that, with the exception of a slight dip in performance for the group that received strategy, there was a gradual increase in reasoning scores across trials for both the group receiving strategy and the group not receiving strategy. This result would be more in keeping with a practice effect rather than a specific effect for strategy.

The current finding regarding the effect of strategy on performance is consistent with Fogg (1992), where all subjects received strategy, and all showed an increase in performance after receiving strategy. However, because of the design, Fogg (1992) was unable to partial out the variance between strategy and practice effect. The current study was specifically designed to examine the variance in
this area and it appears that strategy did not have the predicted effect. Therefore, it seems that practice, not strategy, was responsible for the increase in performance across trials.

The sixth and final hypothesis predicted that, within groups, after completing the final task, subjects would show an increase in self-efficacy regardless of whether they received positive feedback; however, the group receiving feedback would demonstrate higher levels of self-efficacy than the group not receiving feedback. The analysis failed to support the first part of this hypothesis. In point of fact, the orthogonal contrasts of the repeated measures ANOVA showed that, across all subjects, self-efficacy decreased from the first to the second measure, and there was also a second significant drop from Measures 1 and 2 to Measure 3.

It is possible that the novel task that was used in this study was too difficult and, as a result, subjects did not have positive efficacious feelings about their performance. As indicated in Table 4 (p. 31) subjects' scores for the three reasoning trials never reached 50% correct (i.e., Trial 1 = 41.17%, Trial 2 = 40.42%, Trial 3 = 44.50%). The fact that subjects were self-rating their performance at approximately 65% for all trials (See Table 5, p. 32) might indicate that their ratings contained a self-image enhancing component.
The second part of the final hypothesis was supported. The results of the orthogonal contrasts showed that the group receiving positive feedback had a higher level of self-efficacy ($M = 2.75$) than the group not receiving positive feedback ($M = 3.05$). (Note: lower numbers = higher self-efficacy.) This finding is consistent with previous research in the area of "positive mastery experiences" leading to higher levels of self-efficacy (cf. Bandura, 1982; Bandura & Cervone, 1983, 1986; Gist, 1987).

In spite of the significant result which showed that the group receiving positive feedback had higher levels of self-efficacy than the group not receiving positive feedback, one must ask whether a difference of .80 represents a meaningful difference with this type of task? If the mean of 2.75 was merely a step to even higher levels of self-efficacy, then that would indeed be a meaningful step in the right direction. However, this is an empirical question that can only be answered with further research. If the mean of 2.75 for the group receiving positive feedback represented a "ceiling effect," then that is not very meaningful in a computer task. There were various reasons why a more meaningful difference was not realized.

One possible explanation for the result is that the subjects did not perceive that they had engaged in a "mastery experience." Research in this area would suggest that repeated mastery experiences should lead to higher
levels of self-efficacy (Bandura & Cervone, 1983, 1986). As mentioned above, for the group that received the positive feedback, there was a significant increase in performance on the third reasoning trial. This would be considered a positive mastery experience and one would then expect to see a concomitant rise in level of self-efficacy. However, perceptions of mastery were not explored in the current study, but further studies should examine this area.

It is also possible that the sign of the feedback (neutral, positive) impacted subjects' belief in that feedback. It was mentioned earlier that subjects consistently rated their performance over twenty percentage points better than it actually was. It is possible that this rating was a type of self-serving bias which floundered in the face of feedback telling them they had performed even better.

In looking at the responses to the questions regarding subjects' perceptions of the usefulness of the information provided during the study, there were some interesting findings. For example, when asked if they felt feedback was credible, on a scale from 1 (Strongly Agree) to 6 (Strongly Disagree), subjects had a mean score of 2.94. When asked if they felt feedback was accurate, subjects had a mean score of 2.88. (3 = Agree Somewhat.) One might suggest from these scores that subjects were thinking, "If they are not trying to help us, then they must be trying to deceive us."
However, when asked point blank if they thought they were being deceived subjects had a mean response of 3.53. (4 = Disagree Somewhat; 5 = Disagree.) Based on this, it does not appear that they thought they were being deceived.

It is possible that because the feedback was not coupled with individualized strategies for improving performance that it was not perceived as accurate or credible. Ilgen et al. (1979) suggest that individuals have to be able to use the information received in feedback to make positive changes in their behavior before that feedback becomes meaningful. For example, simply telling a worker "You are performing poorly" tells that individual nothing about how he or she needs to change their behavior in order to perform better.

Because the task used in this study was cognitive in nature, there were no objective behavioral indicators of how the individual went about approaching the task. To accurately provide feedback, one would have to consult with each individual to investigate his or her approach to the problem. Based on the results of that talk, individualized strategies could be developed. However, in the current study, this was not done and it is possible that the strategy provided simply did not fit into the repertoire of cognitive processing that the subjects were using to answer the questions.

Still, considering the notion of "sign of feedback"
(Ilgen et al. 1979), it is also possible that not giving consistent feedback to each subject during the course of their participation compromised the results. Fogg (1992) provided the same sign of feedback to each subject after each reasoning trial. Thus, each subject was told consistently that his or her performance was poor, about average, or excellent. Also, they were not asked after each trial to enter a number to represent what they thought was their correct percentage of responses. Rather, they were told consistently how they were performing and asked once at the end of the study to put in a number to represent what they thought their correct percentage of responses was across trials. In Fogg (1992) the figure provided by the subjects closely mirrored the figure they were given in the bogus feedback.

It is possible, then, that consistently telling a person how they are performing without explicitly demanding a self-evaluation leads to the feedback becoming more believable over time. This is especially true of positive feedback and is consistent with the literature on the self-image enhancing properties of feedback (cf. Feather, 1968; Ilgen, 1971; Ilgen & Hamstra, 1972; Shrauger & Rosenberg, 1970). Thus, the results in Fogg (1992) would be expected because the consistency of the feedback made it believable and it was then positively correlated with levels of self-efficacy.
Beyond the design issues, another possible reason that the predictions in the current study were not supported is that of the sample itself. Fogg (1992) used a sample which was composed entirely of undergraduate students and fully one-third or more of them were not psychology majors. The makeup of that subject-pool reduced the possibility that the subjects were "experiment wise." The sample for the current study was composed of virtually all psychology majors, and one-third or more of the pool were in graduate programs in psychology. The current sample was more likely to "see through" the study and respond either positively or negatively to the potential demand characteristics of the study. Further evidence of this was revealed during debriefing sessions when some subjects revealed that they had "figured out" what the study was about. This was due in part to the informed consent form (See Appendix A). Section III of that form states, in part, "...all of the details of the experiment may not have been disclosed prior to the commencement of my participation. If this is the case, all of the details of the study will be explained to me prior to my leaving the testing area." This was a strong indicator that deception of some sort was being used, thus, making the scores on the deception question suspect (See Table 6). Regardless of the directional change in the answers, we believe that the sophistication of the current sample had a negative impact on the results.
Future studies in this area should consider several design issues that would have an impact on the results. First, one must consider whether it is necessary to ask subjects for a self-rating after each performance trial. Bandura and Cervone (1986) suggest that people become motivated to improve when their performance fails to meet either and existing or self-set standard or goal. In the face of a novel task, individuals are inclined to trust the evaluation of an outside source who has "access to some objective predictors of performance attainment" (p. 406). This is in line with Prue and Fairbank's (1981) research which indicates that the feedback "source" has to have both credibility and power.

If a study uses a truly novel task it is unrealistic to expect subjects to be able to accurately evaluate their own performance. Asking them to do so and then providing feedback either real or bogus, but especially the latter, would lead to confusion over the discrepancies between self-ratings and feedback, thus, preventing them from setting realistic goals for future attainment. A better strategy would be to simply provide accurate feedback and allow subjects to strive toward improvement. Another possibility would be to provide feedback only once during the treatment trials; however, if that feedback was not in line with their self-ratings the same result could occur.

In terms of strategy for improving performance, it
would be better to provide the strategy at the beginning of
the session during the sample questions rather than wait
until after the second set of reasoning questions.
Providing the strategy early would make that information
available as subjects developed their own personal method
for solving the problems. Reinforcing the strategy or
simply repeating the strategy between trials would make it
more salient for the subjects and would most likely enhance
their performance on subsequent tasks.

Finally, the ideal sample would be individuals in a
work setting who were familiar with computers and their
capabilities. Short of that, it would be better to have a
sample which was not "experiment wise." Regardless of where
the sample was obtained, we feel that a novel task which was
capable of being quickly understood coupled with strategy
given early in the study along with accurate feedback would
result in a more precise measure of how these components
effect both self-efficacy and computer task performance.
Appendix A

Informed Consent Form
INFORMED CONSENT FORM

I I consent to serve as a subject in the research investigation entitled: Computer-Generated Testing Program for Possible Implementation at CSUSB Psychology Department.

II My participation as a subject in this study is completely voluntary.

III I understand the purpose of this study is to investigate whether a computer can serve as a reliable method of testing, however, all of the details of the experiment may not have been disclosed prior to the commencement of my participation. If this is the case, all of the details of the study will be explained to me prior to my leaving the testing area.

IV I understand my name will NOT be included on the test itself and that the confidentiality of my responses will be maintained at all times.

V I understand that all my questions will be answered before leaving the research study area, that I may refuse to answer any questions presented to me, and that I may withdraw my participation and/or data from the study at any time without penalty or prejudice.

VI I understand that absolutely no experience with computers is required for participation in this research project.

VII I understand that all information collected in this study will be treated as confidential and that no details will be released to anyone outside the research staff.

VIII I understand that I may derive no specific benefit from participation in this study except perhaps from feeling that I have contributed to the development of knowledge in the area of computer testing and extra credit if I so request.

IX I hereby allow the researchers on the project to publish the results of the study in which I am participating with the provision that my name and/or other identifying information be withheld.

X I understand that this study is being conducted by Richard J. Fogg, a graduate student in Industrial/Organizational Psychology, under the
supervision of Matt L. Riggs Ph.D., Psychology Department, CSUSB, (909) 880-5590. I may contact Rich Fogg or Dr. Riggs at any time with my questions, comments, or concerns.

XI I understand that if I have any questions, comments, or concerns about the study or the informed consent process, I may also contact the Department of Psychology Human Subjects Review Board through the Office of the Dean of Graduate Studies (AD-128).

Subject Signature ------- Date Subject Name (Printed)

If you are to receive extra credit for participation in this experiment, please provide the following information: Course to which credit is to be applied (e.g., PSYC-377), and instructor's name.

__________________________  _____________________________
Course                                      Instructor's Name
Appendix B

Demographic Questions
Demographic Questions

1) Enter the last four digits of your Student ID Number: ___________ (Fill in the blank question)

2) Sex: 0 - Male
   1 - Female

3) Age: ___________ (Fill in the blank question)

4) Ethnicity: 1 - African American
   2 - Latino (Hispanic)
   3 - Caucasian
   4 - Asian (Pacific Islander)
   5 - Native American
   6 - Other

5) Have you ever taken a test on a computer?
   0 - No
   1 - Yes

6) I feel comfortable when first encountering a new computer program or application.
   1 - Strongly Agree
   2 - Agree
   3 - Agree Somewhat
   4 - Disagree Somewhat
   5 - Disagree
   6 - Strongly Disagree
   (Note: all subsequent demographic questions use the same response options so they will not be repeated for each question.)

7) When faced with novel situations I achieve successful outcomes.

8) I feel confident my current abilities will be sufficient to ensure success when first encountering a new computer program or application.

9) I feel challenged when first encountering a new computer program or application.

10) I feel compelled to "master the machine" when first encountering a new computer program or application.
Appendix C

Self-Efficacy Scale
Overview of the Self-Efficacy Scale

The original self-efficacy scale (Riggs & Knight, in press) has a total of 31 questions divided into 4 subscales: a 10 item personal efficacy scale, an 8 item personal outcome expectancy scale, a 7 item collective efficacy scale, and a 6 item collective outcome expectancy scale. A Likert-type response format is utilized. Response options include: 1 "Strongly Agree", 2 "Agree", 3 "Agree Somewhat", 4 "Disagree Somewhat", 5 "Disagree", and 6 "Strongly Disagree". Sixteen items on the scale are negatively-worded to create a balance in the scale. Specific instructions are given for each subscale. The four subscales, with instructions, can be seen on the following four pages. The modifications to the personal efficacy scale as used in the current study can be seen on the fifth page following.
Personal Efficacy Scale

Think about your ability to do the tasks required by your job. When answering the following questions, answer in reference to your own personal work skills and ability to perform your job.

1. I have confidence in my ability to do my job.
2. There are some tasks required by my job that I cannot do well.
3. When my performance is poor, it is due to my lack of ability.
4. I doubt my ability to do my job.
5. I have all the skills needed to perform my job very well.
6. Most people in my line of work can do this job better than I can.
7. I am an expert at my job.
8. My future in this job is limited because of my lack of skills.
9. I am very proud of my job skills and abilities.
10. I feel threatened when others watch me work.
Personal Outcome Expectancy Scale

Think about the results of doing your job well OR doing your job poorly. Do important outcomes depend upon how well you perform, or do most job-related outcomes occur whether or not you do a good job?

1. I am well-rewarded for my good work.
2. Doing good work here is not worth the effort.
3. Performing your job well is a sure way to get ahead here.
4. Most of my good work goes unnoticed.
5. Around here, such things as salary and promotions are determined by how well a person does his or her job.
6. My work evaluations are accurate.
7. Good work gets the same results as poor work in this job.
8. I must do a good job in order to get what I want.
Collective Efficacy Scale

Think about the department in which you work. This department may be an office group, a maintenance crew, an academic department, etc. When responding to the following items, answer in reference to this group's work-related ability.

1. The department I work with has above average ability.
2. This department is poor compared to other departments doing similar work.
3. This department is not able to perform as well as it should.
4. The members of this department have excellent job skills.
5. Some members of this department should be fired due to lack of ability.
6. This department is not very effective.
7. Some members in this department cannot do their jobs well.
Collective Outcome Expectancy Scale

Think about the department in which you work. This department may be an office group, a maintenance crew, an academic department, etc. Think about the results of this department doing its job well OR doing its job poorly. Do important outcomes depend upon the department's performance, or do most job-related outcomes occur whether or not the department does a good job? When answering the following questions, answer in reference to your beliefs about your current department.

1. It is important for our group to do good work.
2. Many people benefit when our group does good work.
3. No one would notice if our group did its work poorly.
4. This organization depends heavily upon the quality of work my group does.
5. This organization does not need the work done by this group.
6. My group expects good outcomes when we do good work.
Modified Personal Efficacy Scale

Changes to the original scale are indicated as follows:
Deletions are marked in strikeout,
Changes are marked in underline & italic.

Instructions to subjects:
Think about your ability to do the tasks required while participating in this study. When answering the following questions, answer in reference to your own personal work skills and ability to perform this task.

1) I have confidence in my ability do this task.
   1 - Strongly Agree
   2 - Agree
   3 - Agree Somewhat
   4 - Disagree Somewhat
   5 - Disagree
   6 - Strongly Disagree
   (Note: all questions use the same response options so they will not be repeated for each question.)

2) There are some tasks required by this task that I cannot do well.

3) When my performance is poor, it is due to my lack of ability.

4) I doubt my ability to do this task.

5) I have all the skills needed to perform this task very well.

6) Most people in my line of work could do this task better than I can.

7) I am an expert at this task.

8) My future at this task is limited because of my lack of skills.

9) I am very proud of my job skills and abilities.

10) I feel threatened when others watch me work.
Appendix D

Reasoning Questions with P-Values
<table>
<thead>
<tr>
<th>p-values</th>
<th>Reasoning Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5400</td>
<td>BABY, TODDLER, CHILD : SEEDLING, (A) LEAF, (B) BRANCH, (C) TREE, (D) BARK.</td>
</tr>
<tr>
<td>2.8200</td>
<td>HEAD, NECK, WRIST : BRACELET, (A) ARTIST, (B) SHOULDER, (C) TOE, (D) RING.</td>
</tr>
<tr>
<td>3.6800</td>
<td>BURNT, WELL-DONE, MEDIUM : PINK, (A) RED, (B) BLACK, (C) GRAY, (D) WHITE.</td>
</tr>
<tr>
<td>4.3200</td>
<td>EYE, JAY, KAY : TEA, (A) LAY, (B) SEA, (C) YOU, (D) COFFEE.</td>
</tr>
<tr>
<td>5.4800</td>
<td>CHERRY, ORANGE, GRAPEFRUIT : BASEBALL, (A) NEEDLE, (B) GOLF BALL, (C) FOOTBALL, (D) VOLLEYBALL.</td>
</tr>
<tr>
<td>6.5500</td>
<td>CITY, COUNTY, STATE : PROVINCE, (A) COUNTRY, (B) LORE, (C) TOWN, (D) GOVERNMENT.</td>
</tr>
<tr>
<td>7.3800</td>
<td>ARID, DRY, DAMP : WET, (A) SOAKING, (B) SPEAKING, (C) MOIST, (D) DESERT.</td>
</tr>
<tr>
<td>8.2800</td>
<td>CANTALOUPE, MELON, FRUIT : VEGETABLE, (A) MINERAL, (B) FOOD, (C) LETTUCE, (D) LEMON.</td>
</tr>
<tr>
<td>9.4900</td>
<td>GREAT GRANDFATHER, GRANDFATHER, FATHER : MOTHER, (A) GRANDDAUGHTER, (B) DAUGHTER, (C) GRANDMOTHER, (D) CHILD.</td>
</tr>
<tr>
<td>10.4900</td>
<td>CONCEPTION, BIRTH, MARRIAGE : 23, (A) O, (B) 10, (C) 352, (D) 82.</td>
</tr>
<tr>
<td>11.2800</td>
<td>ROCKET, JET, PROPELLER PLANE : TORTOISE, (A) CHEETAH, (B) HARE, (C) ANT, (D) MARSUPIAL.</td>
</tr>
<tr>
<td>12.4000</td>
<td>RHODE ISLAND, CONNECTICUT, NEBRASKA : GERMANY, (A) CANADA, (B) ALASKA, (C) MONACO, (D) BERLIN.</td>
</tr>
<tr>
<td>13.4200</td>
<td>TABLE OF CONTENTS, INTRODUCTION, CONCLUSION : EPILOGUE, (A) INK, (B) PROLOGUE, (C) PLOT, (D) INDEX.</td>
</tr>
<tr>
<td>14.5100</td>
<td>LOITERING, SMUGGLING, KIDNAPPING : LIFE IMPRISONMENT, (A) MURDER, (B) ELECTRIC CHAIR, (C) DETENTION, (D) CRIME.</td>
</tr>
<tr>
<td>15.7800</td>
<td>ANIMAL, VERTEBRATE, MAMMAL : REPTILE, (A) PIG, (B) BIRD, (C) SNAKE, (D) FROG.</td>
</tr>
<tr>
<td>p-values</td>
<td>Reasoning Items</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td>16..7700</td>
<td>PLAY, ACT, SCENE : BOOK, (A) CHAPTER, (B) LIBRARY, (C) NOVEL, (D) AUTHOR.</td>
</tr>
<tr>
<td>17..3400</td>
<td>ZERO, ONE, TWO : CIRCLE, (A) CUBE, (B) ELLIPSE, (C) SPHERE, (D) ROUND.</td>
</tr>
<tr>
<td>18..1200</td>
<td>4, 3, 2 : CHESS, (A) BRIDGE, (B) PAWN, (C) SOLITAIRE, (D) GAME.</td>
</tr>
<tr>
<td>19..6800</td>
<td>SPRING, SUMMER, FALL : JACKET, (A) DRESS, (B) COAT, (C) WOOL, (D) BOOTS.</td>
</tr>
<tr>
<td>20..3700</td>
<td>CONSERVATIVE, MODERATE, LIBERAL : PROGRESSIVE, (A) POLITICAL, (B) REACTIONARY, (C) DEMOCRATIC, (D) RADICAL.</td>
</tr>
<tr>
<td>21..3700</td>
<td>ANCIENT, MEDIEVAL, RENAISSANCE : CANDLE, (A) LIGHT BULB, (B) ELECTRICITY, (C) SUN, (D) TALLOW.</td>
</tr>
<tr>
<td>22..4300</td>
<td>EIGHTH, QUARTER, HALF : DOUBLE, (A) TRIPLE, (B) QUADRUPLE, (C) NOTHING, (D) FRACTIONAL.</td>
</tr>
<tr>
<td>23..6600</td>
<td>TRUNK, SUITCASE, HANDBAG : PITCHER, (A) CUP, (B) WATER, (C) LITER, (D) AQUARIUM.</td>
</tr>
<tr>
<td>24..1100</td>
<td>VILLAGE, TOWN, CITY : CARACAS, (A) NEW YORK, (B) VENEZUELA, (C) HAMLET, (D) OXFORD.</td>
</tr>
<tr>
<td>25..2900</td>
<td>ENGAGED, MARRIED, SEPARATED : BIRTH, (A) BABY, (B) LIFE, (C) CONCEPTION, (D) DEATH.</td>
</tr>
<tr>
<td>26..8000</td>
<td>MATCH, PENCIL, BROOMSTICK : NEEDLE, (A) WIG, (B) HAYSTACK, (C) SWORD, (D) AUTOMOBILE.</td>
</tr>
<tr>
<td>27..4000</td>
<td>IMPOSSIBLE, UNLIKELY, POSSIBLE : MAYBE, (A) NEVER, (B) YES, (C) UNCERTAIN, (D) UNKNOWN.</td>
</tr>
<tr>
<td>28..5100</td>
<td>HEALTHY, SICK, DYING : HOSPITAL, (A) OFFICE, (B) GRAVEYARD, (C) DUST, (D) OLD AGE HOME.</td>
</tr>
<tr>
<td>29..2600</td>
<td>ALERT, FATIGUED, DROWSY : YAWN, (A) ASLEEP, (B) LAUGH, (C) SIGH, (D) SNORE.</td>
</tr>
<tr>
<td>30..9100</td>
<td>MEASURE, MIX, BAKE : EAT, (A) DIGEST, (B) STOVE, (C) COOL, (D) PAN.</td>
</tr>
<tr>
<td>p-values</td>
<td>Reasoning Items</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td>31..1800</td>
<td>ANTARCTICA, SOUTH AMERICA, CENTRAL AMERICA: MEXICO, (A) CHILE, (B) GUATEMALA, (C) GREENLAND, (D) WASHINGTON D.C.</td>
</tr>
<tr>
<td>32..3400</td>
<td>ONLY CHILD, TWINS, TRIPLETs: TRICYCLE, (A) BICYCLE, (B) CAR, (C) UNICYCLE, (D) ICE SKATES.</td>
</tr>
<tr>
<td>33..8500</td>
<td>FRIEND, GIRL FRIEND, FIANCEE: MISS, (A) MRS., (B) HIT, (C) WIFE, (D) MOTHER.</td>
</tr>
<tr>
<td>34..3800</td>
<td>PLOW, PLANT, HARVEST: GRAIN, (A) MOLECULE, (B) FLOUR, (C) SEED, (D) WHEAT.</td>
</tr>
<tr>
<td>35..1800</td>
<td>A, E I : EYE, (A) SEE, (B) O, (C) OH, (D) EAR.</td>
</tr>
<tr>
<td>36..5400</td>
<td>100%, .75, 1/2, : 3/6, (A) WHOLE, (B) ONE-EIGHTH, (C) .4, (D) 1/4.</td>
</tr>
<tr>
<td>37..4800</td>
<td>DESPISE, DISLIKE, LIKE: GOOD, (A) EVIL, (B) EXCELLENT, (C) BETTER, (D) ADMIRE.</td>
</tr>
<tr>
<td>38..5100</td>
<td>BACH, BEETHOVEN, GERSHWIN: VAN GOGH, (A) PICASSO, (B) MICHELANGELO, (C) SHOSTAKOVICH, (D) REMBRANDT.</td>
</tr>
<tr>
<td>39..2200</td>
<td>ALL, MANY, FEW: SEVERAL, (A) EARLY, (B) NONE, (C) FEW, (D) NUMEROUS.</td>
</tr>
<tr>
<td>40..3500</td>
<td>EINSTEIN, NEWTON, PYTHAGORAS: CHAUCER, (A) SATAN, (B) HEMINGWAY, (C) GALILEO, (D) HOMER.</td>
</tr>
</tbody>
</table>

1 P-values are based on Fogg, 1992.
2 Correct answers are marked in bold text.
Appendix E

Feedback Provided to Subjects
NEUTRAL PERFORMANCE EVALUATION

OVERALL PERFORMANCE: 

___ Excellent

X Average

___ Poor

TESTING SPEED: 

___ Excellent

X Average

___ Poor

TESTING ACCURACY: 

___ Excellent

X Average

___ Poor

AVERAGE PERCENT SCORED CORRECT: ___ 68%___

EVALUATION BASED ON COMPARISON WITH:

___ Previous Performance on Same Task

X Other People Taking Test At This Time

___ Standard Score (Norm) of 100 Previous Test-takers

COMMENTS:

___ Keep up the good work!

X You have a chance to do better on the next set of questions.

___ You can definitely do better on the next set of questions.

When you have finished reading this feedback press the spacebar on the keyboard and the next set of questions will be presented. Please talk to no one else in the room.
### POSITIVE PERFORMANCE EVALUATION

<table>
<thead>
<tr>
<th>OVERALL PERFORMANCE:</th>
<th></th>
<th>Excellent</th>
<th></th>
<th>Average</th>
<th></th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>TESTING SPEED:</td>
<td>X</td>
<td>Excellent</td>
<td></td>
<td>Average</td>
<td></td>
<td>Poor</td>
</tr>
<tr>
<td>TESTING ACCURACY:</td>
<td>X</td>
<td>Excellent</td>
<td></td>
<td>Average</td>
<td></td>
<td>Poor</td>
</tr>
</tbody>
</table>

**AVERAGE PERCENT SCORED CORRECT:** 94%

**EVALUATION BASED ON COMPARISON WITH:**

- Previous Performance on Same Task
- Other People Taking Test At This Time
- Standard Score (Norm) of 100 Previous Test-takers

**COMMENTS:**

- Keep up the good work!
- You have a chance to do better on the next set of questions.
- You can definitely do better on the next set of questions.

When you have finished reading this feedback press the spacebar on the keyboard and the next set of questions will be presented. Please talk to no one else in the room.
Appendix F

Strategy for Improving Performance
TRY THIS STRATEGY FOR IMPROVING YOUR PERFORMANCE:

One of the most frequently encountered kinds of reasoning problems are series completion problems. In these problems, the solver is usually given some terms that form a series of some kind and the task is to complete the series. Series completion problems can involve words or figures. Consider the series completion problem below:

LIE, KNEEL, STAND -- HIGH:

1 - TALL
2 - HIGHER

In this problem the relation to be inferred is one of progressively greater vertical height. This relation needs to be transferred to the term HIGH and then applied so as the yield a correct answer. If HIGHER is not perceived as the ideal answer option, then logic is required to recognize this is better than TALL.

You must infer successive relations between each adjacent pair of three terms. Thus, in the sample, you need to infer, first, the relation between LIE and KNEEL, second, the relation between KNEEL and STAND, and, finally using the same type of relationship, between the target word HIGH and the best answer; in this instance, HIGHER.

Try this strategy on your next set of questions. It will, most probably, help improve your scores.
Appendix G

The Design of the Experiment
## DESIGN OF THE EXPERIMENT

<table>
<thead>
<tr>
<th></th>
<th>GROUP 1</th>
<th>GROUP 2</th>
<th>GROUP 3</th>
<th>GROUP 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Demographic</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>Self-efficacy #1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>Sample questions</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>Reasoning Trial #1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>5</td>
<td>Self-Rating #1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>Feedback #1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>7</td>
<td>Self-efficacy #2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>8</td>
<td>Reasoning Trial #2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>9</td>
<td>Self-Rating #2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>10</td>
<td>Feedback #2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>11</td>
<td>Self-efficacy #3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12</td>
<td>Strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Reasoning Trial #3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>14</td>
<td>Self-Rating #3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>15</td>
<td>Feedback #3</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Self-efficacy #4</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>17</td>
<td>Credibility Ques.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Appendix H

Debriefing Form
DEBRIEFING

Thank you for participating in this study!

The Ethical Principles of Psychologists (American Psychologist, 1981, 36, 633-6380; Principle 9, Research with Human Participants; Paragraph "e," states, in part, "Methodological requirements of a study may make the use of concealment or deception necessary. Before conducting such a study, the investigator has a special responsibility to: (i) determine whether the use of such techniques is justified by the study's prospective scientific, educational, or applied value, and (ii) insure that the participants are provided with sufficient explanation as soon as possible."

The study in which you just participated utilized a degree of deception in three areas:

I) The true name of the study is "The Effects of Feedback and Strategy on Self-Efficacy and Computer Task Performance." The researchers were of the opinion that your knowledge of the true title of the experiment could alter your responses such that outcomes would be altered to a significant degree.

II) All feedback provided to you during the course of the study was bogus. Feedback was based on your participation in one of four different testing groups. All participants received neutral feedback (i.e., scores of 68%) after the first two sets of reasoning questions. Participants in two of the groups received positive feedback (i.e., scores of 94%) after the third set of reasoning questions. Two other groups received no feedback after the third set of reasoning questions. Thus, any similarity between feedback received and your actual performance was purely coincidental. The researchers will statistically evaluate, for example: level of self-efficacy and task performance, any change in level of performance after receiving feedback, any change in level of self-efficacy after receiving feedback, etc.

III) The study has nothing to do with the implementation of a computer testing program within the Psychology Department at CSUSB.

It is hoped that the results of this study will
contribute to knowledge which can be utilized by psychologists within industrial/organizational settings. Specifically, it is hoped that users receiving feedback and strategy via computers will find it to be equally or more useful than feedback which is received from a supervisor in the more traditional verbal or written format.

[Note: The first two pages of the debriefing form were read by the subjects while still in the testing area. These two sheets were retained in the area to preclude the possibility of the true nature of the study being revealed. The third page (next page) was given to the subject to be retained by them.]
As is required by institutional policy and procedure, this study was approved by the Department of Psychology Human Subjects Review Board through the Office of the Dean of Graduate Studies (AD-128). As such, it was determined that the potential benefits of this line of research outweighed the minimal risks which might be visited upon the participants in the study.

However, if you become anxious or upset as a result of participating in this study, please call the Counseling Center (Ext: 5040) in PS-227. They offer free therapy to students. In addition, the local mental health departments provide counseling on a sliding scale so that individuals who have little of no income would pay little or nothing. The number for the San Bernardino Department of Mental Health is (909) 387-7171 and for the Riverside Department of Mental Health the number is (909) 358-4500.

If you have any questions about this research project or would like to find out what the results are when completed please contact:

Matt L. Riggs, Ph.D.
CSUSB, Psychology Department
Office: FO-121
Phone: 880-5590

In order to maintain the confidentiality of this research, it is requested that you discuss the details of your participation and the nature of the study with no other student until the end of the Spring Quarter, at which time data collection should be completed.

Your cooperation in this matter is gratefully appreciated.
APPENDIX I

Questions regarding Subjects Feelings About Study
Introduction:
The next four screens will have questions regarding your feelings about different aspects of this study.

1) I feel that the strategy provided for improving my performance was helpful.
   1 - Strongly Agree
   2 - Agree
   3 - Agree Somewhat
   4 - Disagree Somewhat
   5 - Disagree
   6 - Strongly Disagree
   
   (Note: all questions use the same response options so they will not be repeated for each question.)

2) I feel that the feedback provided to me during this study was credible, or believable.

3) I feel that the feedback provided to me during this study was accurate.

4) I feel that I was deceived during the course of this study.
References


