1989

Effect of sex-appropriate task description on the tendency to self-handicap

Kendell C. Thornton

The University of Montana

Let us know how access to this document benefits you.
Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation
https://scholarworks.umt.edu/etd/5758

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
COPYRIGHT ACT OF 1976

This is an unpublished manuscript in which copyright subsists. Any further reprinting of its contents must be approved by the author.

MANSFIELD LIBRARY
UNIVERSITY OF MONTANA
DATE: 1989
The Effect of Sex-appropriate Task Description on the Tendency to Self-Handicap

By
Kendell C. Thornton
B.S., University of Idaho, 1985

Presented in fulfillment of the requirements for the degree of
Masters of Arts
University of Montana
1989

Approved by

Chair, Board of Examiners

Dean, Graduate School

Date
Oct. 5, 1989
In an effort to uncover some of the determinants of self-handicapping, male and female introductory psychology students chose to listen to either facilitating or interfering music prior to taking a test similar to one they had just previously completed. The music selection served as the dependent measure of self-handicapping. In addition, prior to taking the first test, half of the subjects were led to believe that their sex did better at the test than the other sex, whereas the remainder were led to believe the opposite. Also, half of the subjects were given contingent success feedback after completion of the first test, while the other half received noncontingent success feedback. Results did not reveal any significant main effects or interactions.
INTRODUCTION

Attribution theory is considered by many to be the study of the processes by which people form causal interpretations of events around them (Arkin & Baumgardner, 1985). The self-serving bias in causal attribution refers to the tendency of individuals to attribute successful outcomes to themselves and unsuccessful outcomes to other factors (Arkin & Baumgardner, 1985). Although a failure may have occurred and may not be reversible, denying personal responsibility for unsuccessful outcomes vastly reduces the negative quality of the bad news, and minimizes the implications of the failing outcome. In other words, by attributing an unsuccessful outcome to some extraneous external cause, one can sever the usual link between performance and evaluation. In contrast, by assuming personal responsibility for successful outcomes, the implication of the success in determining one's level of ability is maximized.

The self-serving bias in causal attribution is a fairly clear example of attributional principles aimed at protecting or sustaining one's image. Self-protective strategies can therefore be seen as the use of attributional principles so that only desired inferences about personal qualities can be drawn. For example, "excuse-making" is a
self-protective strategy that is hypothesized to occur when an actor is perceived as being responsible for a negative outcome (Snyder, Ford, & Hunt, 1985). Excuse-making has been defined as any action or explanation that reduces the negative implications of a "bad performance", thereby preserving the individual's positive self-image (Snyder et al., 1985). This maneuver applies the attributional principles retroactively in order to manage one's image.

The present research is concerned with one kind of self-protective strategy, which attempts to control the esteem implications of performance feedback by the choice of setting in which performance occurs. Specifically, the phenomenon of "self-handicapping" was originally defined as "any action or choice of performance setting that enhances the opportunity to externalize (or excuse) failure and to internalize (reasonably accept credit for) success" (Berglas & Jones, 1978, pg. 406). With this type of self-protective strategy the attributional rules must be applied in anticipation of the performance, rather than after some outcome has occurred, or retroactively. Alcohol use may be seen as an example of the self-handicapping strategy. Ability may be viewed as a facilitative cause of successful performance, whereas alcohol is an inhibitory cause. If an individual drinks prior to a performance and subsequently does poorly on that performance, an external attribution can
be made in order to explain the poor performance and maintain self-esteem. However, if the individual does produce a successful performance, the inference that ability is the most likely potential cause of a successful performance is augmented by the presence of alcohol. Therefore, regardless of what the performance outcome is, the self-handicapper has protected or maintained their self-esteem. Other examples of self-handicapping are getting too little sleep or underpreparing before an examination, or in general embracing impediments and plausible performance handicaps (Berglas & Jones, 1978).

Berglas and Jones (1978) designed their experiments to test the proposition linking self-handicapping strategies—in this case the choice of a performance inhibiting drug—to a recent history of noncontingent success. They defined noncontingent success as an operation where success feedback follows a performer's attempt to offer solutions to insoluble problems. More specifically, the performers do not see that the outcome is appropriate to their performance, and they feel that they lack the knowledge or control that success feedback implies. This lack generates the anxious uncertainty that an effective performance cannot be repeated at will. In addition, the experiments investigated whether such self-handicapping is exclusively a maneuver designed to influence the attributions of an
audience or whether it serves self-protective functions as well. To accomplish this, subjects were recruited for an experiment allegedly designed to measure the effects of drugs on intellectual performance. Subjects were either exposed to insoluble problems or to readily soluble problems before being given enthusiastic success feedback. Subjects were then allowed to choose between performance-facilitating and performance-inhibiting drugs with the understanding that the effects would be active during a retest on comparable problems. In half of the cases, the experimenter administering the drugs was clearly aware of the subject's prior intellectual performance, making this a "public" performance; in the remaining cases, he was obviously unaware; thereby making the results of this group's performance "not public". The prediction that subjects would prefer the performance-inhibiting drug only in the noncontingent success condition was clearly confirmed, but this result was true for males only. Also, in line with prediction, the publicity of the male subject's success was not a significant factor in this drug choice. With the female sample, however, there is the hint of an interaction between contingency and publicity, with females tending to follow the drug-choice predictions only in the public conditions.
Since Jones and Berglas' (1978) original work, the topic of self-handicapping has received considerable attention. When evaluative conditions and expectancy instruction encouraged such behavior, subjects found to be Type A (Weidner, 1980), test anxious (Smith, Snyder, & Handelsman, 1982), hypochondriacal (Smith, Snyder, Perkins, 1983), socially anxious (Snyder, Smith, Augelli, & Ingram, 1985), and heavy drinkers (Higgins & Harris, 1988), were shown to self-handicap more than their relevant comparison groups.

Since the original Berglas and Jones (1978) investigation of self-handicapping, many have broadened the original definition of self-handicapping to include verbal claims of a handicap. A close examination of the literature by Leary and Shepperd (1986) led them to conclude that researchers have used the term "self-handicapping" to refer to two different phenomena. Some writers have operationalized the construct in line with Berglas and Jones' original position, as a behavioral strategy, typically described as the adoption of impediments to success in a situation where the person anticipates failure. Thus the person has a self-enhancing explanation for the impending failure and is allowed to maintain self-esteem and the illusion of competence. Other writers have expanded the definition of self-handicapping to include a verbal claim
that a handicap exists which will debilitate one's performance on a task in which ability is implicated. For example, women have been found to claim bad mood as a handicap to an upcoming task when bad mood could serve as an alternative explanation for poor performance (Baumgardner, Lake, & Arkin, 1985). However, unlike the person who engages in a behavior that would be expected to make success more difficult, such as taking a performance-debilitating drug, one who only claims a handicap will not make it more difficult to succeed on the threatening task.

Leary and Shepperd (1986) have suggested that the term "behavioral self-handicapping" be used to refer to the actions of people who construct handicaps that augment nonability attributions for possible failure. Thus, individuals who consume performance-debilitating drugs before evaluative performances, who do not take advantage of factors that would be expected to enhance performance, such as practice, or who elect to be evaluated under difficult circumstances, might be engaging in behavioral self-handicapping. They have further suggested that the label of "self-reported handicap" might then refer to the use of verbal claims that one possesses handicaps that interfere with one's performance. Here, an individual may claim to be ill, anxious, weak, shy, or the victim of traumatic incidents in those instances in which such states ostensibly
excuse poor performance. In this way they are also providing attributionally relevant information that discounts ability attributions for possible failure. The present author agrees with this distinction of the way self-handicapping has been operationalized in the literature so far, and feels that this needs to be taken into account when speaking of the existing findings. So far, however, it is a distinction that has not been well maintained in the literature.

The present research is an attempt to clarify the sex differences which have been shown to exist in the self-handicapping literature. Previous research is inconclusive regarding any sex differences in self-handicapping strategy (Snyder, Ford, & Hunt, 1985). Men have been shown to self-handicap through performance-inhibiting drug selection (Berglas & Jones, 1978; Kolditz & Arkin, 1982), alcohol consumption (Tucker, Vuchinich, & Sobell, 1981; Higgins & Harris 1985; Higgins and Harris 1988), and practicing less when practice was expected to enhance performance (Harris & Snyder, 1986). Although some women have been shown to self-handicap through performance-inhibiting drug selection (Berglas & Jones, 1978) they did not do so at the level that men did, nor was it found to be statistically significant. In the only other study that investigated women's tendency to self-handicap through performance-inhibiting drug
consumption (Bordini, Tucker, Vuchinich, & Rudd, 1986) the self-handicapping hypothesis that noncontingent success would produce relatively greater alcohol consumption was not supported. Regardless of feedback, insoluble test subjects consumed more alcohol than did soluble test subjects. However, since only females were used as subjects, and a social evaluation task which was thought to be more relevant to their self-esteem was employed, it hard to interpret these results. Women have, however, been shown to claim handicaps of test anxiety (Smith et al., 1982), physical symptoms (Smith et al., 1983), self-reports of shyness (Snyder et al., 1985), and traumatic life events (DeGree & Snyder, 1985) when such symptoms could serve as excuses for poor performance.

It appears that Leary and Sheppsards' (1986) conceptual distinction between behavioral self-handicaps and self-reported handicaps directly addresses the sex differences that have been shown to exist so far. Men have more often demonstrated behavioral self-handicapping while women have more often been found to use self-reported self-handicaps. Although men have been shown to self-handicap through self-reports of test anxiety (Greenberg, Pyszczynski, & Paisley, 1985), and self-reports of shyness (Snyder et al., 1985), there does not appear to be any evidence that women actually handicap themselves behaviorally (Harris & Snyder, 1986;
Since the present research is investigating this apparent sex difference, I have chosen to restrict my operational definition of self-handicapping to the original definition (Berglas & Jones, 1978) which appears to be referring to "behavioral self-handicaps".

The sex-appropriateness of the experimental task may be one reason why a sex difference has typically been found. Achievement motivation research (Eccles, Adler, & Meece, 1984) suggests that achievement-related behaviors are influenced by the sex-role stereotyping of experimental tasks and that analytic cognitive tasks are viewed by both sexes as being male sex-typed. Certain achievement contexts (e.g. competition) may also differentially affect the intrinsic motivation of males and females to perform a particular task (McHugh, Frieze, & Hanusa, 1982). Although it is not understood how a task becomes sex-typed, the sex-typing of a task has been successfully manipulated by describing an ambiguous task as involving either masculine or feminine sex-typed skills (McHugh, Fisher, & Frieze, 1982), and by telling subjects that one sex typically performs better than the other on the task (Baucom & Danker-Brown, 1984; Stake, 1976).

From the earliest to the more recent formulations of the self-handicapping concept, an underlying idea is that self-esteem protection is the primary reason one engages in
self-handicapping (Berglas & Jones, 1978; Snyder & Smith, 1982). Generally, the self-handicapper has been thought to have a tenuous or uncertain sense of esteem (Bergals & Jones, 1978; Harris et al., 1986). The certainty or stability of the person's self-esteem, rather than the level of self-worth, may be most relevant to the elicitation of the self-handicapping behavior, at least for male subjects (Harris & Snyder, 1986). Harris and Snyder (1986) found that males with uncertain self-esteem practiced less (self-handicapped more) for the esteem-threatening evaluative task than certain males, and certain or uncertain females. However, although we know that self-esteem is a very important variable, it appears that we do not know all of the elements that may help create a self-esteem threatening situation which will cause a person to engage in self-handicapping.

It has been suggested that the elicitation of self-handicapping behavior requires evaluative stress (Harris et al., 1986). It may be possible to increase the evaluative stress of the experimental task by making the results of the performance public. In the original study by Berglas and Jones (1978), after the initial testing session was completed, the experimenter either was or was not told about the subjects' performance and the publicity of the subject's success was not found to be a significant factor in the drug
choice. This effect was attributed to the more robust male sample, and if the females are looked at alone there appears to be a nonsignificant hint of an interaction between contingency and publicity, with females tending to follow the drug-choice self-handicap only in the public conditions. Tice and Baumeister (1985) found that under public conditions, high self-esteem subjects practiced less (handicapped more) than low self-esteem subjects. They suggest that this was done for self-presentational reasons, presumably in order to provide a protective excuse in case they did not do well at the task. Although Tice and Baumeister used both males and females as subjects they did not look at sex differences, and there does not appear to be any research that addresses the issue of 'publicity' while using both males and females besides Berglas and Jones (1978). Thus, some have concluded that the primary motivational underpinning of self-handicapping is the preservation rather than the enhancement of esteem (Arkin & Baumgardner, 1985; Kolditz & Arkin, 1982), and results suggest that subjects tend to handicap themselves more in the "public" conditions (Baumgardner et al., 1985, Tice & Baumeister, 1985). The original Berglas and Jones (1978) experimental situation was found to have less impact on females than on males. Bordini and his colleagues (Bordini et al., 1986) thought that this sex difference might be a
consequence of the intellectual tasks used in self-handicapping studies concerned with substance abuse, and investigated whether women might self-handicap with alcohol or drug use in response to other evaluative threats. Their experiment employed a social judgment task to investigate whether women who were evaluated as normal drinkers would respond to noncontingent success by consuming relatively greater amounts of alcohol. They found that women faced with performing an insoluble social judgment task consumed more alcohol than those who anticipated performing a soluble version of the task. This effect due to task difficulty held regardless of the feedback manipulation, thus failing to support the self-handicapping hypothesis that noncontingent success would result in the greatest levels of alcohol consumption. They concluded that this lack of support and the negative results obtained by Berglas and Jones (1978) for their women subjects raises the issue of possible sex differences in the use and social approval of the use of alcohol or drugs as self-handicapping strategies. On the other hand, Bordini et al. (1986) proposed that these negative findings may indicate that neither the social task they used nor the intellectual task used by Berglas and Jones engaged the self-esteem of the women subjects to an extent conducive to self-handicapping. Because of the importance of the sex-appropriateness of the experimental
task, the purpose of the present study was to offer a task that was more involving to females. It is possible that the females, perceiving the task of solving analogies to be a male appropriate task, did not feel that the performance outcome was as important to their self-esteem as the males did. This lack of ego involvement may have the resulting effect of decreasing the evaluative stress of the performance and the felt need to engage in self-handicapping. The basic prediction underlying the present experiment is that women will be found to behaviorally self-handicap if they are presented with an experimental situation that is as involving to them, and felt to be as relevant to their self-esteem, as it appears to be for men.

Along with this, a less socially stigmatized self-handicapping option was also provided. Bordini et al. (1986) pointed out that research indicates alcohol abuse carries a greater social stigma for women than for men, and hypothesized that under some conditions, performance failure may be relatively more acceptable to women than risking the apparent negative social consequences of excessive alcohol consumption for women. In order to remain as close as possible to the original Berglas and Jones (1978) studies, I chose to use the typical experimental task of solving analogies and to manipulate the perceived sex-appropriateness of that task. It was hypothesized that
telling women that they were involved in a sex-appropriate task, and that women do better at the task than men, would increase their involvement, or the degree to which they would feel the performance outcome to be a threat to their self-esteem. The anticipated increase in evaluation stress was intended to be sufficient to cause women to engage in behavioral self-handicapping. The self-handicapping measure was tailored after Rhodewalt and Davison's (1986) study in which they had subjects choose whether they wanted to listen to music that would have a "facilitative effect" on their upcoming performance or music that would have an "inhibitory effect" upon their performance. It was assumed that this measure would not carry any undesirable stigma while still providing an adequate self-handicapping option.

The present experiment was designed to test the proposition that women as well as men would be found to behaviorally self-handicap if they were provided with an experimental situation that was sufficiently involving and relevant to their self-esteem. To this end, subjects were recruited for an experiment allegedly designed to measure the effects of music on intellectual performance. Some subjects were exposed to task descriptions that described the task as appropriate for males and others were exposed to task descriptions that described the task as appropriate for females. Half of the subjects were then exposed to
insoluble analogy problems and the rest to readily soluble problems before being given enthusiastic success feedback. Subjects were then allowed to choose which level of music, ranging from extremely facilitating to extremely debilitating, they would like to listen to during a retest on comparable problems. In each of the cases the experimenter was clearly aware of the subjects' performance on the first test, making this a "public" performance. The main prediction was that noncontingent-success subjects, both male and female, who received sex-appropriate task information, would be found to engage in behavioral self-handicapping more than subjects who received sex-inappropriate task information. In line with this prediction, sex-appropriate/noncontingent success subjects were expected to self-handicap more than males and females in the contingent success conditions.
METHOD

Subjects

Subjects were 120 undergraduate males and females (60 of each sex) at the University of Montana, who participated in the experiment in partial fulfillment of their introductory psychology research requirements. Two male experimenters conducted each experimental session.

Overview of the Design and Procedure

The design of the experiment was a 2 (sex typing of task: Appropriate, Inappropriate) x 2 (feedback: Contingent, Noncontingent) x 2 (sex of subject: male, female) between groups design. Subjects participated individually in each experimental session. Each subject was greeted at the lab by an experimenter, who explained that the present study was designed to examine the effects of music on intellectual performance (see Appendixes A & B for complete instructions). Each subject was told they would take "two 20-question tests, each of which contains a randomly selected group of analogies and progressions borrowed from the nationally used 'Cochran's Aptitude Test'". It was further explained that the tests were "designed to discriminate the uppermost levels of intellectual potential" and that they should not be surprised to score no higher than the 60th-70th percentiles. Subjects were told that the
purpose of this procedure was to compare the subject's initial test score, achieved without listening to music, with the test score achieved while listening to music. Additionally, they were told that they would be given feedback regarding their performance at the end of each test.

Sex typing of task manipulation: Each subject was randomly presented with one of two possible sex typing descriptions (see Appendix A). They were told either that "previous research has shown that women score significantly higher than do men on this particular type of test" or that "previous research has shown that men score significantly higher than do women on this particular type of test". In short, each subject was told either that members of their sex score higher than members of the other sex or that members of their sex score lower than members of the other sex. The "sex-appropriate" condition for females then consisted of telling the female subjects that their sex scored higher than the other sex and that the "present research is designed to compare the performance of women like yourself, with men, since men have been found not to do well on these tests, and women have been found to do very well on these tests." The "sex-inappropriate" condition for females stated just the opposite: "men score significantly higher than do women" and "the present research is designed
to compare the performance of women like yourself, with men, since women have been found not to do well on these tests, and men have been found to do very well on these tests". The text for male subjects was essentially the same only tailored to a male audience. The present author conducted a pilot study to test the effect of these instructions, the results of which indicated that the manipulation of perceived sex appropriateness was successful. That is, subjects given sex-appropriate task information anticipated that subjects of their sex would do better at this task than subjects of the other sex and subjects given sex-inappropriate task information anticipated that subjects of their sex would score lower than subjects of the other sex.

After establishing the sex appropriateness of the task, the experimenter then described the nature of the intellectual performance tests, explained what analogy problems are and how to solve them, and allowed the subject to see an example of an analogy problem (see Appendix B).

Contingency manipulation: Subjects were randomly assigned to either contingent- or noncontingent-success conditions. The experimenter administered a 20-question test of analogies (see Appendix C), and their difficulty level plus the subsequently announced success defined the contingency manipulation. Subjects in the noncontingent conditions were given 16 very difficult, relatively
insoluble questions and 4 questions that were easily solved. They were, however, given feedback that they had successfully solved 16 out of 20 problems. The contingent-success condition was achieved by giving subjects 16 relatively easy questions and 4 very difficult, relatively insoluble questions. These subjects were given the same feedback, that they had successfully solved 16 out of 20 problems. Regardless of what condition subjects were in, upon completion of the test they were informed that they "have done exceptionally well" and "Yours was one of the best scores I've seen so far!". The present author conducted a pilot study on his contingent/noncontingent success manipulation, the results of which indicated that the materials and procedures were adequate. That is, contingent-success subjects expressed greater confidence in the accuracy of each problem solution than the noncontingent-success subjects.

In administering the test, the experimenter handed the subjects individual multiple-choice analogy problems printed on index cards. Subjects were given 15 seconds to respond orally to each item as the experimenter coded the answers onto a form. Furthermore, after answering each analogy problem, subjects gave an accuracy rating indicating their certainty that the response was correct, expressed in terms of a percentage between 0 and 100. This follows Berglas and
Jones (1978) and others (Tucker, Vuchinich, & Sobell, 1981; Kolditz & Arkin, 1982), who have found that giving such accuracy estimates seems to enhance the unstable nature of success experienced by subjects who answer insoluble test questions.

After subjects had received their success feedback, an assistant entered the lab and was informed of the subject's performance on the first test. The experimenter left after it was explained that the assistant would be administering the second test. This assistant was blind as to which contingency condition the subject was in, and which sex-type instructions they had received.

Dependent measures: The assistant led each subject into the next room where they were seated at a table with seven tapes and a cassette player with headphones. The tapes were labeled from 1 to 7, with the words "Extremely Facilitating" on tape 1 and "Extremely Debilitating" on tape 7, and tape 4 was marked "Neutral". The main dependent measure was the choice of music to be listened to while taking the second test of ability. This closely follows the procedures used by Rhodewalt and Davison (1986). Subjects were reminded that previous research had indicated that the various forms of music that were being used in the present study were disruptive or facilitative of intellectual performance. Then they were told to select the music they would like to
Listen to while taking the second test. It was stressed that they not be concerned with selecting a particular tape to please the experimenter. Subjects were told that "since we need different subjects to be exposed to each of the levels of music, it really doesn't matter which one you choose, and you can go ahead and select the tape that you would like to listen to while taking the second test."

Subjects then recorded their music choice. After the music selection had been made, subjects were asked to complete a few questionnaires. It was emphasized that the questionnaires were to check their perceptions of their performance on the first test and this experiment as this may affect their performance on the second test. The first questionnaire was a series of manipulation checks disguised as a "mid test questionnaire" (see Appendix D). The second questionnaire was the "Self-description inventory", which is a scale designed to measure an individual's tendency to self-handicap (see Appendix E). The third and last questionnaire was titled "Preliminary Impression Of Intelligence Study" and was designed to probe for suspicion and possible knowledge of the hypothesis. They were told the questionnaires were to check their perceptions of their performance on the first test and this experiment, as this might affect their performance on the second test.
The "Mid test questionnaire" (see Appendix D) asked subjects to indicate on a 9-point scale their answers to 10 questions. For example, they were asked "How well do you feel you did on this test?" (1 = Not well at all, 9 = Extremely well). They were also asked to what extent they thought luck versus ability accounted for their test score; how satisfied they were with their performance; how difficult they felt the intelligence test was; how stressful the test was; how much control they felt they had over their test performance; how difficult they thought the second test would be as compared to the first; what effect they thought the music was going to have on their performance on the second test; how well they expected to do on the second test as compared to the first; and they were asked to estimate how their performance on this test compared to the average college student.

The "Self-description inventory" (see Appendix E) is a measure of an individual's tendency to engage in behaviors that strategically protect self-esteem (Strube, 1986). Correlational analyses with both the long version (Rhodewalt et al., 1984) and the short version (Strube, 1986) have indicated that high self-handicapping was related reliably to high public self-consciousness, high social anxiety, high other-directedness, high depression, and low self-esteem. These relationships were stronger when the short form was
used suggesting the utility of the abbreviated self-handicapping scale for future research (Strube, 1986). However, to date, the scale has seen only limited use, and psychometric data are sparse. It is interesting to note that the self-handicapping scale indicates that high self-handicapping is related to low self-esteem. In a direct test of the hypothesis that level of self-esteem may mediate self-handicapping, Tice and Baumeister (1984) recruited high and low self-esteem subjects to participate in a study in which their abilities would be evaluated. Subjects were told that they would have a chance to practice the task before the evaluation; amount of practice served as the dependent measure of self-handicapping. Their study was also investigating the impression management dimension of self-handicapping and the authors manipulated whether subjects' practice sessions would be observed or not by the experimenter. Although there was no main effect for self-esteem, a Publicity x Self-Esteem interaction indicated that high self-esteem subjects practiced reliably less (self-handicapped more) than did low self-esteem subjects in the public condition, and also less than the high self-esteem subjects in the private condition. The Tice and Baumeister finding that high self-esteem subjects handicapped to a greater degree in public than the low self-esteem subjects is consistent with Harris et al.'s (1985) observation of
self-handicapping performance attributions by high self-esteem subjects. These results contradict, however, the observation of an inverse relation between self-esteem and disposition to self-handicap as theorized by Arkin and Baumgardner (1985) and reported by Jones and Rhodewalt (1982). Because of the differences in these experiments, direct comparisons between the studies are difficult. In light of these inconsistent results, further research is needed. The inclusion of the scale in this research was more motivated by curiosity than by any particular prediction.

The "Preliminary impression of intelligence study" questionnaire (see Appendix F) probed for suspicion about the credibility of the experiment. Subjects were asked how confident they were in the accuracy of the test just completed. They were asked if there was "anything about the intelligence test that would make you doubt its accuracy". Also, they were asked to describe in their own words what they thought the hypothesis of this study was. Last, they were told that often psychology students read in their classes about experiments in which things are not as they seem, and that this occasionally disturbs their natural responses when they are subjects. Then they were asked "While you have been participating in this experiment, have you honestly felt any doubts about any aspects of it? If
so, what were they? Describe in what ways, if any, they have affected your behavior thus far in the experiment". Following this, the subjects were informed that the experiment was over, and they were debriefed. (See Appendix G for debriefing text)
RESULTS

Manipulation Checks

The effectiveness of the test difficulty and feedback manipulations in creating the contingent and noncontingent success conditions was demonstrated through several manipulation check results. Because of the differences in problem solubility that constituted the contingency manipulation, subjects in the contingent-success condition should have registered higher accuracy estimates after each problem solution than noncontingent-success subjects. Each subject's accuracy estimates were totaled and the mean of each subject's accuracy estimates served as the dependent measure. As Table 1 shows, a 2 x 2 x 2 (Sex Typing of Task x Feedback x Sex) analysis of variance (ANOVA) performed on these accuracy estimates indicated that noncontingent-success subjects (M =46.84) indeed expressed less confidence in the accuracy of each problem solution than the contingent-success subjects (M =82.67), F(1,112) =325.23, p<.001.

Insert Table 1 about here

The manipulation check questions from the Mid Test Questionnaire yielded similar results, indicating success in creating the contingency conditions. A t-test for the
difference between two means was used to evaluate the questionnaire items. Since the effectiveness of the two main experimental manipulations was established in the two pilot studies, it was not felt necessary to perform a more elaborate analysis. As anticipated, contingent subjects felt they did better on the intelligence test (M=7.15) than noncontingent subjects (M=5.63), (Appendix D: Question 1), \( t(118) = 4.83, p < .001 \). Also in line with prediction, contingent subjects attributed their performance more to ability (M=6.73) than noncontingent subjects (M=3.92), (Appendix D: Question 2), \( t(118) = 8.04, p < .001 \). In addition, as expected, the contingent subjects were more satisfied with their performance (M=7.17) than noncontingent subjects (M=6.38), (Appendix D: Question 3), \( t(118) = 2.45, p < .05 \); the insoluble problems were seen as more difficult (M=6.88) than the soluble ones (M=3.62), (Appendix D: Question 4), \( t(118) = -11.95, p < .001 \); subjects in the noncontingent condition rated the test as more stressful (M=3.95) than contingent subjects (M=2.35), (Appendix D: Question 5), \( t(118) = -4.91, p < .001 \); contingent subjects claimed to have more control over their test performance (M=7.13) than noncontingent subjects (M=5.30), (Appendix D: Question 9), \( t(118) = 5.94, p < .001 \); surprisingly, contingent subjects did not anticipate the second test to be less difficult (M=6.02) than noncontingent subjects (M=6.05),
(Appendix D: Question 12), $t(118) = -0.14, p = 0.890$; as expected, contingent subjects felt more confidence about how they would do on the second test ($M = 6.30$) than noncontingent subjects ($M = 4.85$), (Appendix D: Question 13), $t(118) = 5.12, p < .001$; and contingent subjects expected to perform better on the second test ($M = 4.87$) than noncontingent subjects ($M = 4.25$), (Appendix D: Question 15), $t(118) = 2.80, p < .05$. Although this was not thought of as a manipulation check when the study was designed, it is interesting to note that when asked "How well do you think the average female (or male) college student would score on this test?" (Appendix D: Questions 6 & 7) contingent subjects felt that the average female would score better ($M = 6.55$) than did noncontingent subjects ($M = 5.67$), $t(118) = 3.25, p < .05$. This may be another indication that contingent success subjects felt the test was easier and so anticipated that it would be easier for the average college student. In line with this observation, contingent subjects also felt the average male would score better ($M = 6.25$) than did noncontingent subjects ($M = 5.53$), $t(118) = 2.70, p < .05$.

Because the Sex-typing of task manipulation was created by telling subjects either that members of their sex score higher than members of the other sex or that members of their sex score lower than members of the other sex, the manipulation check was to see if subjects would correctly
use this information to anticipate how an average female and male college student would score on the test (Appendix D: Questions 6 & 7). Because of the wording of the questions, the ratings on these questions by females who received sex-appropriate task information were expected to differ from females who received sex-inappropriate task information. Similarly, the ratings on these questions by males who received sex-appropriate task information were expected to differ from males who received sex-inappropriate task information. As Tables 2 & 3 show, a 2 x 2 x 2 (Sex Typing of Task x Feedback x Sex) ANOVA was calculated for these two questions. As expected, in response to the question "How well do you think an average female college student would score on this test?", there was a significant interaction between Sex and Sex Typing, $F(1,112)=12.60$, $p<.001$. Females who received sex-appropriate task information anticipated that an average female would score higher ($M=6.27$) than did females who received sex-inappropriate information ($M=5.50$), $t(112)=2.11$, $p<.05$; and as expected, males who received sex-appropriate task information anticipated that females would score less well ($M=5.80$) than did the males who received sex-inappropriate information ($M=6.87$), $t(112)=2.93$, $p<.05$. In response to the question "How well do you think an average male college student would score on this test?", females who received sex-appropriate
task information did not anticipate that an average male would score higher (M=5.73) than did females who received sex-inappropriate information (M=5.63), t(112)=0.27, p>.05. There was a trend for males who received sex-appropriate task information to anticipate that an average male would score higher (M=6.40) than males who received sex-inappropriate information anticipated (M=5.80), t(112)=1.6, .05<p<.10.

In line with the sex typing of task manipulation, the pilot study for this manipulation revealed a possible "macho" trend whereby males told that "women typically do better" reported that they felt more confident, expected the test to be easier, and expected to do better than the males receiving sex-appropriate task information. When asked how well they thought they would do if they knew (males or females) do better (Appendix D, Question 10), each subject was asked about the opposite sex than they were told did better in the instructions. A 2 x 2 x 2 analysis of variance revealed a significant interaction between contingency and sex typing, F(1,112)= 4.57, p<.05. Contingent success/inappropriate subjects did not expect to do better (M=5.33) than contingent success/appropriate
subjects (M=5.53), \( t(112)=0.73, \ p>.05; \) while noncontingent/inappropriate subjects did expect to do better (M=5.53) than noncontingent/appropriate subjects (M=4.90), \( t(112)=2.28, \ p<.05. \) Thus, it appears that the noncontingent/inappropriate subjects were answering in a "macho" manner. This appears to be due to the females in this group. The follow-up question to this asked subjects what reasons they might have for answering the above question the way they did (Appendix D, Question 11). These responses were rated by the experimenters as either "macho", "maybe macho" or "answered as anticipated". An analysis of the content of the answers revealed that most subjects (n=74) did not feel that the sex typing information should make any difference, with only a few subjects actually being rated as "macho" (n=12), and the rest answering as anticipated (n=28), with the exception of one subject being rated as "maybe macho". Therefore, question 11 does not appear to offer support for the "macho" tendency seen in the pilot work.

In line with previous findings by Strube (1986), males and females did not differ in their self-reported self-handicapping tendencies as measured by the Self-Description Inventory, \( F(1,112)=0.43, \ p=0.515. \) In addition, using the median split of the scores on the Self-Description Inventory to indicate high and low self-handicappers, the Self-
Description Inventory was not an accurate predictor of self-handicapping in this study. Pearson product moment correlations were calculated for the relationship between the dependent measure of music selection and the scores on the Self-Description Inventory, \( r(118)=0.026, p>.05 \). Therefore, there was no relationship between an individual's score on the Self-Description Inventory and their music selection.

Self-Handicapping

The principle dependent measure in this investigation was the selection of the type of music each subject wanted to listen to during the second test. Each music selection had a scale value from 1 (extremely facilitating) to 7 (extremely debilitating) with Position 4 designated as a neutral tape. The tape choice constituted the measure of self-handicapping. A 2 x 2 x 2 (Sex Typing of Task x Feedback x Sex) analysis of variance was performed on this measure of self-handicapping and these means are reported in Table 4. There were no significant main effects or interactions to support the hypothesis that sex-appropriate/noncontingent subjects self-handicapped more than the relevant comparison groups, \( F(1,112) =1.23, p=0.291 \).

Insert Table 4 about here
DISCUSSION

The results of the present study did not provide supportive evidence for the hypothesis that noncontingent-success subjects, who received sex-appropriate task information, would be found to engage in behavioral self-handicapping more than subjects who received sex-inappropriate task information. More surprisingly, the present study did not even offer support for the typically robust finding that noncontingent success males self-handicapped the most. In addition, the "Preliminary Impression of Intellectual Study" questionnaire was rated by the experimenters as either "suspicious", "not suspicious" or "maybe". Removing the suspicious subjects data did not significantly change the results. Finally, it appears that this study did provide support for the utility of the Self Description Inventory, as it confirmed that the relevant groups did not differ in their Self-handicapping tendencies.

It is possible that there is a problem with music selection as the measure of self-handicapping. It was observed that subjects would often choose the "Extremely Facilitating" tape and then rate that they expected the music would "impair" their performance on the second test. As soon as this became apparent, subjects were asked during debriefing why this had been their anticipated effect of the
music. The most frequent comment made was "I always find it hard to study while listening to music". It appears that subjects may not have truly believed the labels as to the effects of the music, or they may have simply disregarded the information in preference to their own past experience.

Recently, Shepperd and Arkin (1989) also used Rhodewalt and Davidson's (1984) measure of music selection as the dependent measure of self-handicapping. In a manner similar to the present study, they attempted to manipulate the perceived importance of the task, with the prediction that an important task should elicit greater self-handicapping. However, Sheppard and Arkin accomplished this by telling subjects either that the test they were taking was a reliable and valid predictor of college and career success (high task importance), or that the test was a new test that had yet to be validated or shown useful for any predictive purposes (low task importance). While it can be argued that this is quite different than the present study's sex-typing of task manipulation, it appears that the underlying purpose of both manipulations was to provide subjects with an experimental situation that is sufficiently involving and relevant to their self-esteem.

Sheppard and Arkin's (1989) study also investigated the effects of preexisting environmental handicaps on a subject's tendency to self-handicap. The idea is that the
presence of preexisting handicaps provides an excuse for failure, should it occur, permitting individuals to devote their undivided attention and effort toward performing well. Besides being instructed that they were participating in an experiment investigating the effect of music on performance on an intellectual test, subjects were either told that the test they were taking was a reliable and valid predictor of college and career success or that the test was a new test that had yet to be validated. Subjects in the handicap-present condition were then told that the experiment was also investigating dichotic listening and that subjects would hear a high-pitched intermittent ringing noise transmitted through the left earphone whereas the music they selected would be transmitted through the right earphone. In addition, subjects in the handicap-present condition were informed that the purpose of the study was to determine whether the inhibiting music would compound the detrimental effect of the ringing noise and whether the facilitating music would cancel out the detrimental effect of the ringing noise. Subjects in the handicap-absent condition were told nothing of the dichotic-listening task.

Subjects were shown sample test items, and told they would complete a test composed of similar items. They were then told they would have 10 minutes to complete the test and were provided with a test booklet and answer sheet.
Next, subjects were instructed to select the cassette they wished to listen to while taking the test. After the music had been selected, it was made apparent that the cassette player did not work. The experimenter inspected the cassette player as if to try to get it to work. Finally, the experimenter announced that the experiment would have to be cancelled. Then, as if acting on a second thought, the experimenter asked subjects to complete the postexperimental questionnaire as some of the data would be useful. The experimenter also stated that some of the questions might not make sense since the subject had not taken the test, and the experimenter simply requested that the subject respond as though they had taken the test.

As predicted, individuals receiving high importance instructions were more likely to handicap their performance in anticipation of taking a test, but this occurred only when no other handicap was present. It is interesting to note that these results were obtained in anticipation of taking the test, and without providing subjects with feedback about their ability as the typical self-handicapping study has done. Although it does break away from the original paradigm of Berglas and Jones (1978), it may be that this is an even better paradigm to measure an individual's tendency to self-handicap as it eliminates the difficult manipulation of contingent and noncontingent
success, and allows the subject to react more in accord with their past experience. However, it is not readily apparent why Sheppard and Arkin were able to obtain significant results using music selection as the measure of self-handicapping and the present author was not.
REFERENCES


APPENDIX A

MUSIC AND INTELLECTUAL PERFORMANCE

Today, you will be participating in a project designed to investigate the facilitating and inhibiting effects of music on intellectual performance.

You will be taking two 20-question tests, each of which contains a randomly selected group of analogies and progressions borrowed from the nationally used "Cochran's Aptitude Test".

These tests are "designed to discriminate the uppermost levels of intellectual potential", so you should not be surprised to score no higher than the 60th-70th percentiles. You will be given feedback regarding your performance at the end of each test.

You will be taking two tests. The first test will be administered without music in order to establish a baseline with which to compare your performance on the second test. The second test will be administered while listening to music.

Previous research has shown that women/men score significantly higher than do men/women on this particular type of test. Previous research also indicates that the various forms of music that we will be using in the present study either disrupted or facilitated intellectual performance. The present research is designed to compare the performance of people like yourself, who have been found to do well on these tests, with people who have been found not to do as well on these tests.

The two tests you will be taking have been specifically constructed for this type of research. They are multiple choice analogy problems in which the questions cover a variety of subjects, and each test has been standardized to allow for a 'test-retest' comparison of an individual's scores.

According to its publisher, "The Psychological Corporation", the test is a "high-level mental ability test which requires the solution of a series of intellectual problems stated in the form of analogies".

In other words, the test you will take is an aptitude test. In addition, this test also presupposes a firm grounding in several academic disciplines, including literature, social studies, mathematics, and science; so far as this is the case, this test is also an intelligence test.

Now, let me tell you something about analogies and what you can expect on this intelligence test.
APPENDIX B

Instructions: Stalking the Analogy

The Cochran analogy has four parts, or terms, three of which are given, with the remaining term to be selected from four choices.

"When correctly completed"- that is, when you select the proper choice among the four- "a relationship which exists between two of the terms can be found between the remaining two terms".

This kind of item permits the use of content from many subject matter areas in exercises requiring the perception of a variety of relationships.

To hunt the analogy, you must first understand the meaning of the three terms that are given in the item. Then, you must determine or "perceive" a relationship between two of the terms that are given. The third step is then to choose among the four choices so that the same relationship will exist between the second pair of terms as exists between the first pair of terms.

When this process has taken place, you will have solved the analogy presented in that item.

Example: Light:Dark :: Pleasure: (a.picnic, b.day, c.pain, d.night). "Light is to dark", "as pleasure is to pain". You will find it rather easy to accept that light is to dark as pleasure is to pain, since it is obvious that light is the opposite of dark and pleasure is the opposite of pain.

Are there any questions?

Ok, I will hand you individual questions printed on index cards. You will have 15 seconds to respond orally to each item. Furthermore, after you give your answer to each item, I want you to give a subjective probability estimate of its accuracy, expressed in terms of a percentage between 0 and 100.

Example: "I'm pretty sure, but not positive of its accuracy, I'd say I'm 85% sure".

Or: "I'm positive I got this one correct, I'd say 100% accurate".

Or: "I'm not at all sure, it was mostly a guess, I'd say I'm 10% sure".

Are there any questions?
APPENDIX C

Analogy test items for Contingent Success Condition. (Note: Items 3,4,10 and 15 are the insoluble items)

1. Where : (a. place, b. person, c. thing, d. lost) :: When : Time

2. Frost : Flowers :: Jealousy : (a. love, b. fire, c. rain, d. tree)

3. Lunkhead : Dolt :: Recant : (a. bottle, b. disavow, c. secant, d. profess)


5. Aesop : Fable :: Grimm : (a. sonata, b. essay, c. fairy tale, d. epigram)

6. Tiger : Claw :: (a. tree, b. wolf, c. rose, d. lion) : Thorn

7. Hexagon : 6 :: (a. square, b. sphere, c. pentagon, d. line) : 5

8. Square : Cube :: (a. plane, b. math, c. circle, d. geometry) : Sphere

9. Igloo : Eskimo :: Tepee : (a. hut, b. Indian, c. home, d. wigwam)

10. Ephemeral : Durable :: (a. heated, b. fleeting, c. solid, d. wilting) : Lasting

11. Scales : Fish :: Quills : (a. coat, b. wood, c. porcupine, d. fur

12. Death : Birth :: Last : (a. end, b. first, c. morgue, d. time

13. School : Fish :: Swarm : (a. birds, b. bees, c. buffaloes, d. tigers)

14. Eating : Fat :: Dieting : (a. thin, b. candy, c. carbohydrates, d. protein)
15. Words : Long :: (a. nouns, b. worth, c. actions, 
   d. speech) : Fellow

16. Chicken : Egg :: Plant :: (a. tree, b. bush, 
   c. seed, d. grass)

17. Train : Whistle :: Car :: (a. siren, b. horn, 
   c. song, d. propeller)

18. Gun : Holster :: Sword :: (a. plowshare, b. scabbard, 
   c. relic, d. soldier)

19. Gasoline : Car :: Blood :: (a. body, b. nose, 
   c. knife, d. transfusion)

20. Colt :: (a. revolver, b. dog, c. horse, d. ranch) 
   :: Fawn :: Deer

Analogy test items for Noncontingent Success Condition.
(Note: Items 1, 2, 10 and 15 are the easily solved items)

1. Where :: (a. place, b. person, c. thing, d. lost) 
   :: When :: Time

2. Frost : Flowers :: Jealousy :: (a. love, b. fire, 
   c. rain, d. tree

3. Lunkhead : Dolt :: Recant :: (a. bottle, b. disavow, 
   c. secant, d. profess

   Hamilton, c. Adams, d. Monroe)

5. Tropical : Mental :: (a. luxurious, b. fevered, 
   c. colorful, d. pictorial) :: Aliment

6. Ailurophile :: (a. cats, b. dogs, c. monkeys, d. snakes) 
   :: Philatelist :: Stamps

7. Quatrain : Sonnet :: Haiku :: (a. tanka, b. fusee, 
   c. herpes, d. jihad)

8. Mellifluous : Cacophonous :: Encomium (a. approbation, 
   b. denunciation, c. sophistry, d. palaver)
9. Vacuum : Arraign :: (a. epilogue, b. halvah, c. cleaner, d. radii) : Ebullient

10. Gasoline : Car :: Blood :: (a. body, b. nose, c. knife, d. transfusion)


12. Knoll : Plumb :: (a. lead, b. hill, c. ptarmigan, d. tubing) : Autumn

13. Rubber : Contract :: (a. eraser, b. void, c. expand, d. synthetic) : Doubleton


15. Colt :: (a. revolver, b. dog, c. horse, d. ranch) :: Fawn : Deer

16. Dovecote : Pigeons :: Hutch :: (a. spiders, b. butterflies, c. rabbits, d. snakes)

17. Ether : Nitrous oxide :: Sodium Pentothal :: (a. novocaine, b. angostura bark, c. bakelite, d. sodium bicarbonate)

18. Proboscis : Prognathous :: Nose :: (a. foot, b. ear, c. lips, d. jaw)

19. Ephemeral : Durable :: (a. heated, b. fleeting, c. solid, d. wilting) : Lasting

20. Words : Long :: (a. nouns, b. worth, c. actions, d. speech) : Fellow
APPENDIX D
MID TEST QUESTIONNAIRE

Please indicate (by circling the appropriate number) the degree to which you would answer each of the following questions.

1. How well do you feel you did on the intelligence test you just took?

   Not well 1...2...3...4...5...6...7...8...9 Extremely well
   at all

2. To what extent did luck versus ability account for your test score?

   Factors 1...2...3...4...5...6...7...8...9 Factors beyond my control, such as luck
   under my control such as ability

3. How satisfied are you with your performance?

   Not at all satisfied 1...2...3...4...5...6...7...8...9 Extremely satisfied

4. How difficult did you feel the intelligence test was?

   Extremely easy 1...2...3...4...5...6...7...8...9 Extremely difficult

5. How stressful was this test?

   Not at all stressful 1...2...3...4...5...6...7...8...9 Extremely stressful

6. How well do you think an average female college student would score on this test?

   Not well 1...2...3...4...5...6...7...8...9 Extremely well
7. How well do you think an average male college student would score on this test?

Not well 1...2...3...4...5...6...7...8...9 Extremely well

8. Estimate how your performance on this test compares to the average college student.

Much worse 1...2...3...4...5...6...7...8...9 Much better than the average college student

9. How much control did you feel you had over your test performance?

No control at all 1...2...3...4...5...6...7...8...9 Very much control

10. How well do you think you would do on a test of this nature if you knew that women (or men) typically did better than men (or women)?

Much worse 1...2...3...4...5...6...7...8...9 Much better

11. What reasons might you have for answering question #10 the way you did?

12. How difficult do you think the second test will be as compared to the first?

Less difficult 1...2...3...4...5...6...7...8...9 More difficult

---

1 There were two versions of the Mid Test Questionnaire, with question 10 being the only difference. Subjects received the version that asked about the opposite sex than they were told did better in the instructions.
13. People differ on their level of confidence concerning intelligence tests. How confident are you regarding how well you are going to do on the second test?

Not at all confident 1...2...3...4...5...6...7...8...9 Very confident

14. What effect do you think the music will have on your performance on the second test?

Very great facilitation 1...2...3...4...5...6...7...8...9 Very great impairment

15. How well do you expect to do on the second test as compared to the first?

Much worse 1...2...3...4...5...6...7...8...9 Much better
APPENDIX E

SELF-DESCRIPTION INVENTORY

Please indicate (by checking the appropriate category) the degree to which you agree with each of the following statements as a description of the kind of person you think you are most of the time. Code for responses: AVM = agree very much, APM = agree pretty much, AL = agree a little, DL = disagree a little, DPM = disagree pretty much, and DVM = disagree very much.

1. I tend to make excuses when I do something wrong.

2. I tend to put things off until the last moment.

3. I suppose I feel "under the weather" more often than most people.

4. I always try to do my best, no matter what.

5. I am easily distracted by noises or my own daydreaming when I try to read.

6. I try not to get too intensely involved in competitive activities so it won't hurt too much if I lose or do poorly.

7. I would do a lot better if I tried harder.

8. I sometimes enjoy being mildly ill for a day or two.
9. I tend to rationalize when I don't live up to others' expectations.

/_______/_______/_______/_______/_______/
AVM    APM    AL    DL    DPM    DVM

10. I overindulge in food and drink more often than I should.

/_______/_______/_______/_______/_______/
AVM    APM    AL    DL    DPM    DVM
APPENDIX F

Preliminary Impression Of Intelligence Study

Name_________________  
Sex_____  

1. Was there anything about the intelligence test that would make you doubt it's accuracy?  

2. Describe in your own words what you think the hypothesis of this study is (that is, what is the purpose of the study?)?  

3. Often psychology students read in their classes about experiments in which things are not as they seem, and this occasionally disturbs their natural responses when they are subjects. While you have been participating in this experiment, have you honestly felt any doubts about any aspects of it? If so, what were they? Describe in what ways, if any, they have affected your behavior thus far in the experiment.
APPENDIX G

Debriefing Text

At this point, you have completed this experiment. You will actually not be taking the second test and I would like to explain enough about this research and the reasons for the deception so that you will not be left with any ill effects or misperceptions. We are trying to see if we can get people to perceive a task as being something that they will do well at, by telling them that past research has shown that members of their sex have been found to score higher than the other sex. Some subjects were told that members of their sex have been found to score lower than the other sex. Actually past research has not shown this, and I have simply created this idea as a part of my study. So, you should disregard the notion that one sex does better at this type of task than the other.

I am also interested in studying some of the things people do if they are given success feedback after an intelligence test when they are not sure of their success, and may not feel able to duplicate this success. The test you took was not an intelligence test. It may resemble one as it was constructed so that it would appear to be legitimate. After taking the test all subjects are given the same success feedback, no matter what their answers were. In fact, we did not even take your answers into consideration before we gave you the success feedback. So, please disregard the notion that you just took an intelligence test, and the results which are not in any way reflective of you intelligence.

Deception in psychology research is often used to enable the experimenter to control the subjects' perceptions of the experiment. In this case, we used deception to ensure that subjects felt either that their sex had been found to score higher or lower than the other sex, and to ensure that subjects felt either unsure of their success or confident of their success. The rest of what we did- the music selection you made- was the principle measure we were interested in after you experienced this success feedback.

Due to the nature of this experiment, and its use of deception, you are asked to keep your experience and knowledge about it confidential so that other potential subjects will not be informed prior to taking part in this study. Can I count on you not to talk about the experiment to others for the next couple of months while I finish my data collection?

Are there any questions?
### TABLE 1

Means and Standard Deviations of the accuracy estimates for the problems on the first test.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Males (n=60)</th>
<th>Females (n=60)</th>
<th>Total (n=120)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Contingent Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>85.05/6.8</td>
<td>83.40/7.4</td>
<td>84.22/7.0</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>83.76/6.0</td>
<td>78.49/11.3</td>
<td>81.12/9.3</td>
</tr>
<tr>
<td>B. Noncontingent Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>49.71/10.7</td>
<td>40.32/14.4</td>
<td>45.02/13.3</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>52.77/13.3</td>
<td>44.55/13.7</td>
<td>48.66/13.9</td>
</tr>
</tbody>
</table>

Analysis of variance for accuracy estimates.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sex</td>
<td>1</td>
<td>1129.1</td>
<td>1129.1</td>
<td>9.53</td>
<td>0.003</td>
</tr>
<tr>
<td>2. SexType</td>
<td>1</td>
<td>2.2</td>
<td>2.2</td>
<td>0.02</td>
<td>0.891</td>
</tr>
<tr>
<td>3. Contingency</td>
<td>1</td>
<td>38524.4</td>
<td>38524.4</td>
<td>325.23</td>
<td>0.000</td>
</tr>
<tr>
<td>4. 1x2</td>
<td>1</td>
<td>11.3</td>
<td>11.3</td>
<td>0.10</td>
<td>0.758</td>
</tr>
<tr>
<td>5. 1x3</td>
<td>1</td>
<td>214.7</td>
<td>214.7</td>
<td>1.81</td>
<td>0.181</td>
</tr>
<tr>
<td>6. 2x3</td>
<td>1</td>
<td>341.0</td>
<td>341.0</td>
<td>2.88</td>
<td>0.093</td>
</tr>
<tr>
<td>7. 1x2x3</td>
<td>1</td>
<td>43.1</td>
<td>43.1</td>
<td>0.36</td>
<td>0.548</td>
</tr>
<tr>
<td>Error</td>
<td>112</td>
<td>13266.7</td>
<td>118.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>53532.6</td>
<td>449.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 2

**Question 6:** How well do you think an average female college student would score on this test? (1=Not well; 9=Extremely well)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Males</th>
<th>Females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean/S.D. (n=60)</td>
<td>M/S.D. (n=60)</td>
<td>M/S.D. (n=120)</td>
</tr>
<tr>
<td><strong>A. Contingent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>6.60/1.6</td>
<td>6.47/0.8</td>
<td>6.53/1.3</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>7.33/1.3</td>
<td>5.80/1.3</td>
<td>6.57/1.5</td>
</tr>
<tr>
<td><strong>B. Noncontingent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>5.00/1.8</td>
<td>6.07/1.4</td>
<td>5.53/1.7</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>6.40/1.1</td>
<td>5.20/1.6</td>
<td>5.80/1.5</td>
</tr>
</tbody>
</table>

**Analysis of variance for question 6.**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sex</td>
<td>1</td>
<td>6.075</td>
<td>6.075</td>
<td>3.04</td>
<td>0.084</td>
</tr>
<tr>
<td>2. Sextype</td>
<td>1</td>
<td>0.675</td>
<td>0.675</td>
<td>0.34</td>
<td>0.562</td>
</tr>
<tr>
<td>3. Contingency</td>
<td>1</td>
<td>23.408</td>
<td>23.408</td>
<td>11.70</td>
<td>0.001</td>
</tr>
<tr>
<td>4. 1x2</td>
<td>1</td>
<td>25.208</td>
<td>25.208</td>
<td>12.60</td>
<td>0.001</td>
</tr>
<tr>
<td>5. 1x3</td>
<td>1</td>
<td>4.408</td>
<td>4.408</td>
<td>2.20</td>
<td>0.148</td>
</tr>
<tr>
<td>6. 2x3</td>
<td>1</td>
<td>0.408</td>
<td>0.408</td>
<td>0.20</td>
<td>0.652</td>
</tr>
<tr>
<td>7. 1x2x3</td>
<td>1</td>
<td>1.408</td>
<td>1.408</td>
<td>0.70</td>
<td>0.403</td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td>112</td>
<td>224.000</td>
<td>2.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>119</td>
<td>285.592</td>
<td>2.399</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## TABLE 3

Means and Standard Deviations for Question 7.

### Question 7: How well do you think an average male college student would score on this test?

(1=Not well; 9=Extremely well)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Males  (n=30)</th>
<th>Females (n=30)</th>
<th>Total  (n=120)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean/S.D.</td>
<td>M/S.D.</td>
<td>M/S.D.</td>
</tr>
<tr>
<td><strong>A. Contingent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>6.73/1.3</td>
<td>5.93/1.4</td>
<td>6.33/1.4</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>6.40/1.3</td>
<td>5.93/1.3</td>
<td>6.17/1.3</td>
</tr>
<tr>
<td><strong>B. Noncontingent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>6.07/1.3</td>
<td>5.53/1.8</td>
<td>5.80/1.6</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>5.20/1.3</td>
<td>5.33/1.8</td>
<td>5.27/1.5</td>
</tr>
</tbody>
</table>

### Analysis of variance for question 7.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sex</td>
<td>1</td>
<td>5.208</td>
<td>5.208</td>
<td>2.46</td>
<td>.119</td>
</tr>
<tr>
<td>2. Sextype</td>
<td>1</td>
<td>3.675</td>
<td>3.675</td>
<td>1.74</td>
<td>.190</td>
</tr>
<tr>
<td>3. Contingency</td>
<td>1</td>
<td>15.400</td>
<td>15.400</td>
<td>7.29</td>
<td>.008</td>
</tr>
<tr>
<td>4. 1x2</td>
<td>1</td>
<td>1.875</td>
<td>1.875</td>
<td>0.89</td>
<td>.348</td>
</tr>
<tr>
<td>5. 1x3</td>
<td>1</td>
<td>1.408</td>
<td>1.408</td>
<td>0.67</td>
<td>.416</td>
</tr>
<tr>
<td>6. 2x3</td>
<td>1</td>
<td>1.008</td>
<td>1.008</td>
<td>0.48</td>
<td>.491</td>
</tr>
<tr>
<td>7. 1x2x3</td>
<td>1</td>
<td>0.208</td>
<td>0.208</td>
<td>0.10</td>
<td>.754</td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td>112</td>
<td>236.800</td>
<td>2.114</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>119</td>
<td>265.592</td>
<td>2.232</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 4

Means and Standard Deviations for the Dependent measure: Music selection (1=Facilitating music; 7=Debilitating music)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Males (n=60)</th>
<th>Females (n=60)</th>
<th>Total (n=120)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean/S.D</td>
<td>M/S.D</td>
<td>M/S.D</td>
</tr>
<tr>
<td>A. Contingent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>4.67/1.8</td>
<td>3.87/1.6</td>
<td>4.27/1.7</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>4.93/2.0</td>
<td>3.73/1.4</td>
<td>4.33/1.8</td>
</tr>
<tr>
<td>B. Noncontingent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Appropriate</td>
<td>4.40/2.4</td>
<td>5.00/1.8</td>
<td>4.70/2.1</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>4.47/1.9</td>
<td>3.67/1.7</td>
<td>4.07/1.8</td>
</tr>
</tbody>
</table>

Analysis of variance for Music selection.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sex</td>
<td>1</td>
<td>9.075</td>
<td>9.075</td>
<td>2.66</td>
<td>0.105</td>
</tr>
<tr>
<td>2. Sextype</td>
<td>1</td>
<td>2.408</td>
<td>2.408</td>
<td>0.71</td>
<td>0.402</td>
</tr>
<tr>
<td>3. Contingency</td>
<td>1</td>
<td>0.208</td>
<td>0.208</td>
<td>0.06</td>
<td>0.805</td>
</tr>
<tr>
<td>4. 1x2</td>
<td>1</td>
<td>6.075</td>
<td>6.075</td>
<td>1.78</td>
<td>0.184</td>
</tr>
<tr>
<td>5. 1x3</td>
<td>1</td>
<td>6.075</td>
<td>6.075</td>
<td>1.78</td>
<td>0.184</td>
</tr>
<tr>
<td>6. 2x3</td>
<td>1</td>
<td>3.675</td>
<td>3.675</td>
<td>1.08</td>
<td>0.301</td>
</tr>
<tr>
<td>7. 1x2x3</td>
<td>1</td>
<td>1.875</td>
<td>1.875</td>
<td>0.55</td>
<td>0.460</td>
</tr>
<tr>
<td>Error</td>
<td>112</td>
<td>381.600</td>
<td>3.407</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>410.992</td>
<td>3.454</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

56