INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Hyman, Cynthia

REWARD CONTINGENCY, STANDARDS, AND INTRINSIC MOTIVATION

City University of New York PH.D. 1985

REWARD CONTINGENCY, STANDARDS, AND INTRINSIC MOTIVATION

by

CYNTHIA HYMAN

A dissertation submitted to the Graduate Faculty in Educational Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1985

This manuscript has been read and accepted for the Graduate Faculty in Educational Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

4-29-85
date

Barry J. Zimmerman
Chairperson of Examining Committee

4/29/85
date

Shirley Feldmann
Executive Officer

Professor M. Fish

Professor D. Rindskopf
Supervisory Committee

The City University of New York

Abstract

REWARD CONTINGENCY, STANDARDS, AND
INTRINSIC MOTIVATION
by
Cynthia Hyman

Adviser: Professor Barry Zimmerman

The present study focused on the effect of differential presentation of rewards on children's subsequent perceptions of competency and subsequent levels of intrinsic interest in an enjoyable task. Reward contingency and standards were orthogonally manipulated. Rewards were presented either as contingent upon successful task performance (performance-contingent rewards), contingent upon task engagement (task-contingent rewards), or not presented (no rewards). Standards, defined as objective, non-normative and achievable, were presented prior to task engagement, or not presented. Ninety-six children were presented with the target task--a puzzle-like drawing activity predetermined to be of high initial interest. The reward, an award certificate personalized with the name and photo of the child, was presented according to

experimental condition. Perceptions of competency were measured by questionnaires. Level of subsequent intrinsic interest was defined as time on target task during a free play period following the experimental manipulation. Data were analyzed by a prior contrasts, analysis of variance, and Pearson product-moment-correlations.

As hypothesized results indicated that both perceived personal competency and level of intrinsic interest were significantly higher when rewards were presented as performance-contingent and with standards provided, than when they were presented as task-contingent and without standards. Manipulation of reward contingency and standards yielded no statistically significant main effects on subsequent level of intrinsic interest. The pattern of results were interpreted as supporting the contention that reward contingency cannot independently predict subsequent levels of intrinsic interest in a target task. Presentation of standards as opposed to no standards, and presentation of a performance-contingent reward as opposed to a task-contingent reward had significant positive effects on perceived personal competence. The data were interpreted as consistent with the contention that perceived personal competence is a key intervening

variable which is related to intrinsic interest in a task. The results were discussed in terms of implications for classroom application, including behavior modification programs.

Acknowledgements

I am very happy to finally have an opportunity to formally express my gratitude to the many people who have helped me throughout the long years of hard work.

I would like to thank the entire faculty of the Educational Psychology Department at the Graduate Center. They gave generously of their time, and provided me with an excellent education.

I must personally acknowledge my indebtedness to Professor Barry Zimmerman, my dissertation chairperson. He guided my in all aspects of my research with scrupulous erudition coupled with kindness and understanding.

The love and support of my very special friends never had to be asked for. It was always there. It was always appreciated.

My greatest thanks must be reserved for my magnanimous family. Without their unconditional love the hard times would have been impossible. Thank you Mel, for always being there when I needed you. Thank you Donna, Marc, and Jay, for your selflessness and independence.

With gratitude, I dedicate this work to my parents who provided me with love and understanding, guided me in righteous paths, and instilled in me an unquenchable thirst for knowledge.

TABLE OF CONTENTS

Chapter 1

Introduction

It is accepted practice to encourage students to learn their lessons by rewarding them for appropriate behavior and achievements. Verbal praise, gold stars, and good grades have been used to motivate generations of students. Currently, there is an increasing trend to institute behavior modification programs where rewards are contingently presented. The particular nature of the rewards used, the conditions under which they are given, and the extent to which they are utilized varies depending upon the individual school and teacher. But, in one way or another, students are often tangibly rewarded for the display of appropriate learning behavior.

Although it is generally recognized that children come to school with a natural desire to learn, most teachers depend upon some form of external motivation in their efforts to foster learning. This is usually based on the premise that the child's natural interest in the activity will be enhanced by the addition of extrinsic rewards. Some theorists (Bem, 1967; deCharms, 1968) have

challenged that assumption by contending that the presentation of external rewards may in some instances decrease, rather than increase, motivation.

DeCharms' (1968) theory of personal causation maintains that man desires to be the "origin" of his behavior, and not a "pawn" to external sources. Man prefers activities in which he perceives himself as the locus of causality for his own behavior (intrinsically motivated), to those in which he perceives the locus of causality for his behavior as external to himself (extrinsically motivated). deCharms (1968) contended that when external rewards are offered for behavior which is intrinsically motivated, the person perceives that the locus of control has shifted from himself to the external reward. He therefore concluded that the introduction of external rewards for behavior which was intrinsically motivated may serve to decrease rather than increase the total level of motivation.

Bem (1967) in a reinterpretation of Festinger's (1957) cognitive dissonance studies proposed an attribution and self-perception theory. He maintains that "an individual's belief and attitude statements and the beliefs and attitudes that an outside observer would attribute to him are often functionally similar in that both sets of statements are partial 'instances' from the same evidence" (p. 186). Motivation toward a task is

assumed to be based upon the individual's perception of the locus of causality for his task. When there are observable external rewards to behavior, the person attributes causality for his task performance as due to these rewards; when the observable incentives are insufficient to account for the performance of the task, causality is attributed to the self.

Consequently, when external motivation is unnecessarily high and therefore psychologically oversufficient, a person might infer that his/her actions were motivated by the desire for extrinsic reward, and not by intrinsic interest in the activity itself. Lepper, Greene, and Nisbett (1973) have termed this the "overjustification" hypothesis. It implies that a persons intrinsic interest in a goal may be undermined by inducing him to perform that activity in order to obtain some external reinforcement.

The implications of the overjustification hypothesis are significant for educational practices. It implies that external rewards may sometimes be detrimental to the student, for when these rewards are withdrawn, intrinsic motivation may be diminished. Since the ultimate goal of education is to teach and encourage students to learn on their own initiative, independently of the teacher and classroom, the use of external rewards may ultimately be self-defeating.

Attribution theory is sometimes interpreted as a direct challenge to the use of behavior modification programs in the classroom. The common concern, to cognitive attributionist theorists and operant behaviorists, is whether external rewards can sometimes serve to reduce subsequent intrinsic interest in a target task when those rewards are no longer forthcoming. The goal is to delineate those conditions which differentially determine the subsequent effects of rewards on intrinsic interest.

Experimental research, focusing on the effects of extrinsic rewards on subsequent intrinsic interest, has been prolific. The research however, has become difficult to interpret, with individual projects focusing on different variables, and sometimes reporting contradictory results. There have been multiple attempts to explain the discrepancies in experimental findings, but none have been definitive. Interpretations are more often theoretically and hypothetically based rather than experimentally confirmed.

A promising distinction involves the concept of competency. In 1975 Deci recognized that a person's intrinsic interest is enhanced when he feels competent at a task. He revised his cognitive evaluation theory to include three propositions. The first is basically the original attribution theory indicating that a decrease in

intrinsic motivation will occur "under certain circumstances, when someone receives extrinsic rewards for engaging in intrinsically motivated activities" (p. 139). Second, he hypothesized that if a reward serves to enhance a person's "feelings of competence and self-determination" (p. 141) intrinsic motivation for that task will increase. According to Deci's new model, rewards convey competency when receipt of the reward provides information about the subject's successful performance of the target task. Third, he stated that rewards have a controlling and informational aspect, and depending on whether the rewards are primarily controlling or informational, intrinsic interest will decrease or increase (Deci, 1975).

Arkes (1978) proposed a more parsimonious version of Deci's (1975) theory, one which eliminated the need to determine a priori the controlling or informational aspects of the reward. He proposed that, "Any reward that causes feelings of competence will insulate the behavior from the overjustification effect" (Arkes, 1978, p. 204).

Experimental research ensued which was designed to distinguish between the effects of rewards which were contingent on successful performance, as opposed to those which were contingent on task participation regardless of outcome (Bogiano & Ruble, 1979; Enzle & Ross, 1978;

Karniol & Ross, 1977). Their common hypothesis was that rewards contingent on successful performance engender competency and consequently do not result in loss of intrinsic interest. This hypothesis was supported in the three studies.

It remains unexplained, however, why earlier research (Deci, 1971; Greene & Lepper, 1974; Lepper & Greene, 1975) had found that rewards contingent on successful performance did lead to a significant loss of intrinsic interest in the target task.

There is a significant and consistent difference between those studies which found performance contingent rewards to lead to loss of intrinsic interest, and those that did not. The earlier studies on performance contingent reward did not include standards; those later studies designed their projects to include specific attainable standards. When objective standards were given to the subject prior to task engagement, and when these standards were attainable by the subjects, contingent rewards consistently did not decrease intrinsic interest on the target task.

Conflicting experimental findings as to the effects of rewards on intrinsic interest can be theoretically resolved based on differential utilization of reward contingency and standards. When a reward is presented, contingent upon attaining specified objective attainable

standards, subsequent intrinsic interest will be enhanced. When a reward is presented, contingent upon task engagement, and without any indication of standards, subsequent intrinsic interest will be diminished.

The present study is designed to incorporate the experimental manipulation of reward contingency and of standards. Since the definition of the terms "contingent", "non-contingent", "task-contingent" and "performance-contingent" are variously defined in studies to date, a common definition was adopted in this paper. A task-contingent reward is defined as an expected reward presented contingently for task engagement. Performance-contingent reward is defined as an expected reward presented contingently for successful task performance.

The effects of these manipulations on subsequent intrinsic interest on the target task will be statistically measured. Intrinsic interest will be measured behaviorally, by time on target task during free play and conditions of no reward. In addition, the hypothesized intervening changes in perceived feelings of competency, will be measured by self-report measures.

The contributions of this study to educational psychology are both theoretical and practical. On a theoretical level, the study ought resolve seemingly contradictory experimental findings by delineating those

conditions which differentially determine subsequent effects of rewards on intrinsic interest. Practically, it should indicate how extrinsic rewards can be utilized in group settings, without fear of loss of subsequent intrinsic interest on the target task. It should focus the attention of educators on the importance of engendering competency feelings whenever extrinsic rewards are utilized, and demonstrate ways in which this can be implemented. It ought to delineate those conditions under which extrinsic rewards can be detrimental to subsequent intrinsic interest, so that such methods may be avoided in future educational practice.

Chapter 2

Review of Related Research

Early Research Studies

The attributional theories of Bem (1967) and deCharms (1968) predicted that the presentation of rewards may in some instances serve to decrease motivation. Specifically, they maintained that under conditions where there was high level of intrinsic interest in a task, the introduction of rewards would serve to decrease subsequent motivation in that task. Starting in the 1970's experimental research was conducted to directly test hypotheses concerning the possible negative effects of extrinsic rewards.

The earliest research commonly cited as supporting the theory of negative effects of extrinsic rewards are Harlow's monkey experiments (David, Settlage, & Harlow, 1950; Harlow, Harlow, & Meyer, 1950). These studies were designed to investigate monkey's performance on mechanical puzzles when there was no extrinsic incentive. Harlow was primarily interested in studying learning motivated by a manipulation drive. A secondary purpose was to investigate the effects of a subsequent introduction of a food reward. The focus of interest, herein, is on the effects of the introduction of food, an extrinsic reward, on the monkeys subsequent performance.

In both of these experiments the introduction of food had an immediate disruptive rather than facilitative effect on the monkeys' performance. Therefore, Harlow's studies are frequently cited as evidence that "extrinsic rewards can undermine the intrinsic interest that one has in an activity" (Weiner, 1974, p. 99).

Harlow, Harlow, & Meyer (1950), in an experiment designed to demonstrate a manipulation drive, gave four rhesus monkeys twelve days of experience in manipulating and opening a mechanical puzzle whose solution did not lead to any apparent reward. The control group had the identical puzzle placed in their home cages, but these were left open so that the monkeys could not practice opening them. On the thirteenth day, both groups were observed for five, five-minute periods. The experimental monkeys were significantly more efficient in solving the puzzles, both in terms of total number of solutions, solutions attained in sixty seconds, and ratio of correct to incorrect responses. The authors posited a manipulation drive, which they maintain is primary, to account for the learning and maintenance of the puzzle performance.

In order to discount a secondary reinforcement explanation, Harlow et al. (1950) introduced food as a reward, since food was the only extrinsic incentive previously provided for these subjects. On the fourteenth day the apparatus for the experimental monkeys

was changed to include a food-well. The hasp of the puzzle, when in locked position, covered the food-well, and food was accessible only by solving the puzzle. While the animal watched, the food-well was baited with a raisin, and the mechanical puzzle locked into place. The monkey was then allowed into the test cage and observed for a five-minute period.

The authors reported that the "introduction of food rewards seriously disrupted the efficient puzzle solution which they had repeatedly demonstrated previously" (Harlow, et al., 1950, p. 231). The monkeys made more errors and had no success in solving the puzzle. The disruptive effect of the introduction of food on the monkeys tendency to solve the puzzle is clearly evident. Harlow maintained that this supported his explanation of a primary manipulation drive and weakened an alternative explanation which attributed the monkeys activity to secondary reinforcement.

The disruption in puzzle solving activity after the introduction of food, is often interpreted as support for the hypothesis that extrinsic rewards decrease intrinsic motivation. However, careful analysis of the study does not lead to confirmation of this contention. Deci (1971), a supporter of the attributional hypothesis, acknowledges that the evidence is not clear-cut. First, the number of subjects used was very small, with only

four in each group. Secondly, periods of observation were very short, only five minutes, so that it is possible that the disruption was only a short term effect. There is, in addition, another factor which is of overriding importance. The authors clearly outlined an adaptation procedure which was initiated in an attempt to accustom the animals to obtain food in the experimental cage situation. This adaptation procedure has been overlooked in reviews of the study, yet it is crucial to the interpretation of the results.

For adaptation purposes all the monkeys were brought into the laboratory on the fourth to thirteenth day of the experiment. For the first five days each animal was allowed to feed from two food-wells in the test cage. On each of the next five days, the experimenter baited one food-well and covered both wells with identical objects. The monkeys were then allowed to enter the cage, displace the object, and take the raisin. This adaptation procedure was run independently of the experimental manipulations, and was completed on the thirteenth day, before the introduction of food into the experimental paradigm.

The authors designed this procedure for adaptation purposes, but it is herein maintained that it also served as a learning period for the monkeys. Since each animal was given twenty trials, and each trial ended only after

the monkey was "allowed to displace the object and take the raisin" (p. 229), it is apparent that the monkeys learned to move the obstacle to obtain the food. In addition, the animals' behavior on the fourteenth day confirms that learning did take place during this adaptation period. On the fourteenth day, the monkeys were placed into an almost identical situation, the only difference being that the obstacle covering the food-well, the hasp, could be removed only by solving the puzzle. The monkeys, however, did not respond to the puzzle. Rather they tried to do what they had learned on the 100 previous adaptation trials, they tried to move the hasp.

The learning effects of this adaptation period are also clearly, if unawaringly, reported by Harlow et al. (1950). They relate that subsequent to the introduction of food on the fourteenth day, the monkeys failed to solve the puzzle, and acted in a manner very different from that previously:

> The difference in the kind of errors is also very striking. In the initial 10 tests without food reward, the monkeys never approached the problem by touching the hasp first; in the subsequent tests, all three monkeys always erred by attacking (literally) the hasp first. (p. 231)

During the adaptation period the monkeys learned to remove an obstacle from the food-well, and this is precisely what they attempted to do when placed in a

similar situation, even though it did in fact disrupt puzzle solving behavior.

The monkeys' behavior on the fourteenth day has usually been interpreted as supporting the hypothesis that the introduction of an extrinsic reward can adversely affect a behavior which is intrinsically motivated. The introduction of food, an extrinsic reward, disrupted puzzle solving behavior, which was intrinsically motivated. The evidence, on face value, appears clear-cut. However, the design in the Harlow et al. (1950) study was, as outlined, contaminated by the adaptation procedure. Therefore no conclusions should be drawn from this study about the possible negative effects of extrinsic rewards on intrinsically motivated behavior.

In a second experiment designed to demonstrate the existence of a primary manipulation drive, Davis, Settlage, and Harlow (1950) modified the procedure in several ways. Performance was measured over a longer period of time, brain-operated as well as normal macaque monkeys were used as subjects, and a more complicated puzzle was utilized.

Sixteen monkeys, only four of which were normal, were randomly divided into a food and no-food group. The experiment was divided into three periods of twelve days each. During the first and third periods the puzzle was presented to all subjects without any food incentive.

For the second period, the procedure for the no-food group was unchanged, but each animal in the food group had a raisin placed under the hasp of his puzzle at each assembly. These changes in procedure led to observable differences in the monkeys' behavior and to some different conclusions.

The primary hypothesis, positing the existence of a manipulation drive, is again interpreted as having been supported. The no-food group showed an increase in puzzle solving behavior from period I to period II, and a decrease from period II to III. (The increase is interpreted as due to learning, and the decrease as probably due to satiation.) The interest herein is on the food group, which in this experiment showed increased puzzle solving behavior throughout the three periods. The authors note some initial disruption in puzzle solving activity when food was introduced, but report that this was followed by a marked improvement in performance which continued into the final no-food period.

The results of the Davis et al. (1950) experiment challenge rather than support the hypothesis that external rewards decrease intrinsic motivation. The procedural differences from the Harlow et al. (1950) study may account for the difference in results. In this experiment the monkeys were observed for twelve days rather than for five minute periods of one day. Deci's

(1971) contention that the negative effects brought about by the extrinsic rewards may only be short-term, was experimentally supported. Secondly, there was no report of an adaptation period, so that contamination of results can be eliminated. Most important is the fact that in this study, the raisin actually served as a reinforcer. For twelve consecutive days the raisin was placed beneath the hasp, and was obtainable only after a successful solving of the puzzle. As operant theorists would predict, puzzle solving behavior increased. In this study, an extrinsic reward served to increase target behavior; there are no negative effects attributable to extrinsic rewards.

Although the Davis et al. (1950) study demonstrated that external rewards can serve to increase intrinsic motivation, there are several weaknesses to this study that preclude the drawing of conclusions concerning possible effects of extrinsic rewards. First and foremost to be considered is the subject population. Of the sixteen monkeys who served as subjects only four were anatomically normal, whereas the other twelve monkeys had been subject to varying degrees of brain surgery. Yet despite the fact that "operative conditions yielded the greatest difference in the experiment" (p. 309), the results were reported in global terms of food and no-food groups. Moreover, despite random assignment, the food

and no-food groups differed substantially during the first period, before the introduction of food or any other experimental manipulations. Considering the fact that the groups were not balanced from the outset, and that only two out of eight subjects in the experimental groups were anatomically whole, any generalizable conclusions drawn form this study must be considered as extremely tentative. No generalizeable conclusions as to the positive or negative effects of extrinsic rewards should be drawn from Harlow's studies.

Edward Deci (1971, 1972a, 1972b) conducting research in the laboratory and in the field, was the first to specifically design his experiments to directly test hypotheses concerning the negative effects of extrinsic rewards. Intrinsic motivation was defined as being either innate or learned and was inferred when an activity was performed and there was no apparent reward except for the activity itself. Extrinsic motivation was inferred when the activity was performed and was reinforced by external rewards. Deci (1971) presented his main hypothesis as follows:

> If a person is engaged in some activity for reasons of intrinsic motivation, and if he begins to receive the external reward, money, for performing the activity, the degree to which he is intrinsically motivated to perform the activity decreases. (p. 108)

Deci (1971, 1972a, 1972b) attempted to provide empirical support to the above hypothesis using human, rather than animal, subjects. He suggested that human data might be more fruitfully viewed using a cognitive approach.

Deci (1971, 1972a, 1972b) focused attention on the comparative effects of money and verbal approval, as external rewards, and hypothesized that different rewards may have different effects, depending on a persons cognitive evaluation of them. Deci (1971) maintained "that money and closely related tangible rewards" lead to "cognitive reevaluation" of an activity from one that is intrinsically motivating to one that "should probably not be done with pay". Verbal rewards, which reflect social approval, were interpreted as "less likely to be perceived by the person as controlling his behavior," so there is no stimulus to initiate the process of reevaluation, and no loss of intrinsic interest was predicted. (p. 107)

In Deci's (1971) first study, twenty-four college students participated in three separate one hour sessions of puzzle solving. The Soma puzzle, commercially produced by Parker Brothers, was selected to be the task "because it seemed that most college students would be intrinsically motivated to do it" (p. 108). During each of the three sessions the subjects were asked to

reproduce, using the puzzle pieces, four configurations which were drawn on paper. The difference between the experimental and control groups was that the experimental group received "contingent" reinforcement, one dollar for each puzzle solved within the time limit during the second session; the control group received no payment. Additionally, during the third session experimental subjects were informed that they would not be paid since "there was only enough money available to pay them for one of the sessions" (p. 109).

In order to get a measure of motivation, all subjects were led to believe that they were alone and unobserved for a short period during each of the sessions. On a contrived pretext (that the experimenter had to determine which Soma puzzles to present next to the subject), the experimenter left the room for an eight minute period in the middle of each of the three sessions. During this time subjects were told they could do whatever they liked, i.e., read magazines left in the room, work on the puzzle configurations whose drawings were left on the table, or stare around the room. Subjects were observed by the experimenter through a one-way mirror and the primary measure of motivation was the amount of free time spent by the subjects working on the Soma puzzle configurations.

Comparisons were then made between the rewarded experimental group and the non-rewarded control group, on the amount of free choice time they spent working on the puzzles in the first and third session. This was interpreted as reflective of changes in level of intrinsic motivation, due to presentation or external rewards. It was a measure of intrinsic motivation after the reward was explicitly removed. Deci (1971) computed a simple gain score comparison (time 3 - Time 1) for the two groups and found that the difference was statistically significant at the .10 level. Although this did not reach the customary .05 level of significance he maintained that it did lend support to his hypothesis.

In a second laboratory experiment, Deci (1971) replicated the first experiment with a different student population, with one major difference--he used verbal rather than monetary rewards. He maintained that verbal rewards would have the opposite effect of monetary rewards, i.e., that verbal reinforcement serves to enhance intrinsic motivation, but that monetary and tangible reinforcement decreases intrinsic motivation. He hypothesized that verbal rather than monetary reward will lead to an increase, rather than a decrease in the level of intrinsic motivation. Again comparing the experimental groups' intrinsic motivation in session III

minus its intrinsic motivation in session I, to that of the control groups, he reported a significant positive relationship at the .05 level of significance. This was interpreted as supportive of his hypothesis that verbal reinforcement serves to increase intrinsic motivation.

Deci's (1971) studies attempted to measure experimentally the effects of external rewards on intrinsic motivation. The external rewards were either monetary, verbal or absent. The rewards were presented contingent upon reaching an objective standard--matching the cubes to a given configuration. The achievement of these standards by the subjects was not reported, but it was clear that at least some of the subjects failed parts of the task. Deci (1971) reported that, "any subject who was unable to do a configuration within the 13 minutes allowed was shown the solution" (p. 109). The extent of the failure rate was not reported, and no inference was made by Deci (1971) as to possible effects of this intermittent failure.

The dependent measure of intrinsic motivation was inferred. Deci (1971) used the Soma puzzle task on the "assumption" that the task would be interesting. In a subsequent review (1972a) Deci reported that, "pilot testing substantiated that the puzzles were indeed intrinsically motivating" (p. 220). Specifics were not reported.

In an attempt to verify empirically the assumption that the subjects were intrinsically interested in the task, at the end of each session all of the subjects were asked to rate on a nine-point scale the degree to which they found the Soma puzzle task interesting and enjoyable. The session averages in both experiments were sufficiently high, ranging from 7.25 to 8.25, and Deci accepted this as substantiating the intrinsic motivating quality of the task. However, it must be noted that these averages did not significantly vary between the experimental or control groups, or among the different experimental sessions. Therefore, if the questionnaire scores are accepted as a valid indication of level of intrinsic motivation then it must be simultaneously acknowledged that the manipulation of extrinsic motivation did not significantly affect the level of intrinsic motivation in either of the two studies. Deci did not attempt to explain this inconsistency in reasoning.

Deci (1971) presented his experiments on monetary and verbal rewards as parallel, except for the type of reward presented, but there were two other differences in the experimental paradigm that should be noted. Subjects who failed to match the configuration and cubes were differentially treated in the monetary and verbal conditions. In the money reward manipulation, subjects

who failed were not rewarded. In the verbal reward manipulation, subjects who failed were told that "it was one of the most difficult of all the configurations used and that most people were unable to solve it, so he had not done badly" (p. 112). In the monetary condition the subject, in effect, lost a dollar that could have been earned. Since all subjects were shown that solving the task was possible, the implicit message was lack of competency. In the verbal condition, there was a direct attempt to minimize feelings of failure. Secondly, there was a difference in session III, which may also have affected the results. Subjects in the monetary condition were told that they were not being paid because of lack of funds. This may have conveyed a message that this type of "work" merits payment. The verbal group received no directions as to withdrawal of rewards. These differences were unplanned, and their possible effects have not been acknowledged, to my knowledge.

The major hypothesis received only marginal statistical support ($p < .10$). In actuality, the results were less substantial than indicated, since differences between the experimental and control group were very much dependent upon large initial differences between groups. In the first study, the initial mean difference between the experimental and control groups in time spent working on the puzzle during the free choice

period was 34.3 seconds; in the second study it was 112.8 seconds. Due to these large initial differences, gain scores do not adequately reflect the observed behavior patterns of the subjects. In both studies the total time spent on the task during all three free periods negated rather than supported the respective hypotheses. When monetary rewards were given the total time spent was increased; when verbal rewards were used, the total time spent on task decreased. Similarly, in considering only the first and third periods of the study, when no rewards were given, there was only a slight difference in total time spent on the task--a difference of 9 seconds for the first study, and 48.2 seconds, in the opposite direction predicted, for the second study. The results were clearly biased by the large initial differences between the groups, and no statistically valid conclusions can be drawn from these experiments.

Deci (1971) also conducted a field experiment, in which he again hypothesized that extrinsic rewards would decrease intrinsic motivation. The setting was a college newspaper office, where the task was to write headlines according to prescribed rules. For this study the ten weeks were broken into three periods of four, three, and three weeks. The only difference between the experimental and control group was that the experimental group received fifty cents per headline written during

the second period. The rate of performance in writing the headlines was used as a measure of motivation. Data were analyzed by a comparison of gain scores, and the results were marginally significant at the .10 level. The results of this field study were of questionable validity because they were based on a population of only four subjects per cell, suffered from a high drop-out ratio, and reached only marginal statistical significance.

The methodology used by Deci (1971) in this field study, raises an interesting theoretical question. He used rate of performance as a measure of intrinsic motivation and assumed that "the more quickly one performed, the more highly motivated he was to do this task" (p. 110). Yet, it is also possible to assume that motivation can be directed toward the quality of performance rather than toward speed. In fact, it is plausible that motivation toward better performance could lead to slower rather than faster performance. This confusion between quality and quantity of performance as a measure of intrinsic motivation remains unresolved, and a possible source of misinterpretation in this and subsequent experiments.

Deci (1972a) reported on a follow-up study where reward was presented in a manner termed as "non-contingent"; competency on the task was not required

for payment. Experimental subjects were told that they would be paid two dollars for participating in the experiments, "since there were research funds available for paying subjects" (p. 225). They were informed that they would receive the money at the end of the experiment.

The experimental paradigm differed somewhat from Deci's 1971 studies. There was only one session in the 1972a study rather than three. Following the experimental manipulation when the experimenter left the room, subjects in the 1972a study were told that the problem solving was completed rather than incomplete. The target task, Soma puzzles, and reward, money, were the same. The dependent measure of intrinsic interest was again time on task during free play.

Deci (1972a) reported no significant differences between the experimental and control group. He concluded that non-contingent payment did not affect intrinsic motivation. He interpreted this as consistent with cognitive evaluation theory, reasoning that subjects are more likely to perceive that the rewards are the reason for their performance when money is contingent than when performance is not directly tied to the money.

The manipulation of the independent variable, termed by Deci (1972a) as "non-contingent" reward, is not congruent to "task-contingency" as defined in the

introduction to this paper. Deci promised a reward for participating in the experiment, rather than for task engagement on the target Soma task. Subjects may have cognitively linked the reward to participation in experimental research, rather than directly to the Soma task. Standards of performance were objectively stated, i.e., matching the Some cubes to the drawn configurations. Their attainability, and degree of success by the subjects, was not reported.

The dependent measure of intrinsic motivation, time on target task during free play was taken prior to the receipt of any funds, since the reward was given at the end of the experiment. Deci (1972a) measured the effects of a promised reward on intrinsic motivation, rather than of a reward actually received.

Deci (1972b) attempted to clarify differences between a promised reward and received rewards and to account for differences in the level of intrinsic motivation that are dependent upon timing of payment, by linking it to Adams' (1963) theory of inequity. Adams (1963) maintained that when a subject is given payment for an activity, and feels substantially overpayed, he will experience inequity, and if allowed, will continue to work for a short period for no pay in order to restore equity.

Deci (1972b) utilized the one session paradigm, with Soma puzzles as the target task, and money as the reward. Reward was presented contingent upon attainment of specified standards, successful completion of each configuration. The success rate of the subjects was not reported. The timing of the reward was manipulated--given before free play, after free play, or not at all. Verbal reinforcement was either added or omitted. Intrinsic interest was measured by time on task during free play. The results of the questionnaire were not reported.

Deci (1972b) posited three hypotheses:

> Hypothesis I: When a person is rewarded with money for performing an intrinsically motivated activity, his intrinsic motivation will decrease.
> Hypothesis II: When a person is rewarded with verbal reinforcements for performing an intrinsically motivated activity, his intrinsic motivation will increase.
> Hypothesis III: However, when a person who is performing an intrinsically motivated activity feels inequitably overpayed, he will increase his performance (i.e., make additional inputs) to restore equity. (p. 115)

These hypotheses were tested using a 3 X 2 X 2 (Money and Timing X Verbal Reinforcement X Sex) analysis of variance, with the amount of time the subjects spent working on the task during the free choice period as the dependent measure. The experimental manipulation of money, with three levels (money after, no money, and money before) was highly significant ($p < .005$). Since

the ordering of the three levels was as predicted, Deci (1972b) interpreted these findings as supporting Hypotheses I and III.

Hypothesis II, that verbal reinforcement increases intrinsic motivation, was supported only for males, and not for both sexes. Deci (1972b) attributed these sex differences to the fact that the experimenter in this study was a very "personable and attractive male graduate student" such that females all received interpersonal reinforcement (even in the no verbal reinforcement group). However, this account does not explain why intrinsic motivation was substantially decreased when females in the no money group were given verbal reinforcement.

Deci (1972b) interpreted his results as reflecting "timing of payment". He did not deal with the differences between promised rewards and received rewards, and did not attempt to reconcile differences between theses results and those obtained in his 1971 study. In 1971 Deci reported a negative effect for receipt of reward; in 1972b he cited an increase in intrinsic interest due to receipt of a reward.

Deci (1972b) did not interpret his results in terms of behavior modification programs. However, they indicated that intrinsic interest could be increased by the presentation of a monetary reward. Specifically, in

the "money-before" condition, students who had received a reward for their successful performance spent more time on target task during free play, than students in the non-rewarded condition. Performance contingent rewards resulted in an increase in subsequent intrinsic interest.

The experimental research conducted by Deci (1971, 1972a, 1972b) did not provide consistent and statistically supportable evidence concerning the effects of tangible rewards on intrinsic motivation. His studies did stimulate an abundance of experimental research focusing on the effects of tangible rewards on subsequent levels of intrinsic interest.

Kruglanski, Friedman, and Zeevi (1971) focused their study of the negative effects of extrinsic incentive on several qualitative aspects of task performance. They examined a range of dependent variables including such aspects of task performance as recall and creativity, as well as measures of task motivation.

Kruglanski et al. (1971) studied 32 teen-aged subjects, all of whom had volunteered to participate in a survey on youth movements at Tel-Aviv University. The group was randomly divided into a "no-incentive" and "incentive" group. The incentive group was told that in appreciation for their participation they would be given a guided tour of the Tel-Aviv University psychology

department. The no-incentive group received no such promise.

Following allocation of the subjects to their groups, two measures of creativity, and two measures of recall were administered. After completion of these tasks the subjects answered a questionnaire concerning their enjoyment of the tasks and their willingness to participate in similar projects in the future. Kruglanski et al. (1971) reported that, as predicted, recall and creativity performance were found to be significantly higher in the no-incentive group than in the external incentive group ($p < .01$). Rated enjoyment and willingness to participate in future projects was also greater in the no-incentive compared to the external incentive condition ($p < .05$, $p < .09$, respectively). Kruglanski, et al. (1971) interpreted these results as supporting their hypothesis that higher quality of task performance and motivation would occur in the absence of extrinsic incentives.

The 1971 study by Kruglanski, Friedman, and Zeevi differed from that of Deci (1971, 1972a, 1972b) both methodologically and theoretically. Kruglanski et al. (1971) described their reward as an "incentive", it was not monetary, but rather was a tour of the university psychology department. Furthermore, the tour was only

promised during the study, and the date was set for their visit only after the study was completed.

Kruglanski et al. (1971) were not concerned with initial level of intrinsic motivation. They maintained that, "the more positive attitude toward the task observed under the low incentives results from attributing causality to one's self ... as opposed to external goals under the high incentive" (p. 606). They hypothesized a main effect due to extrinsic motivation and not an interaction between extrinsic motivation and level of intrinsic interest.

Kruglanski et al. (1971) described their incentive as "non-contingent", because it was given for volunteering to participate in the study. It was not similar to task-contingent reward as defined in this paper since even task engagement was not a condition for receiving the reward. It differed from the contingent reward used by Deci (1971, 1972b), where task completion was required.

The dependent measures used by Kruglanski and his colleagues were unique. They dealt with immediate effects, during the reinforced period, rather than following the rewarded activity. They focused on changes in quality of task performance during the reinforced period. The measures of intrinsic interest were

self-report in nature. Unlike prior studies there were no behavioral dependent measures of intrinsic interest.

Kruglanski, Alon & Lewis (1972) conducted a study of retrospective misattribution and task enjoyment. The major thesis of this research was that children would attribute causality for their behavior to a prize, even though in reality it had not caused the behavior, and had in fact been introduced only after the behavior had occurred. Elementary school students from four classrooms voluntarily participated in a series of team competitions. Classes were randomly assigned to a "prize" or "no prize" condition. In the prize condition, a prize was announced and distributed to each participant of the winning team after the completion of the games. Yet instructions to the students inferred that a prize had been previously promised. "'As we said before, members of the winning team will be awarded special prizes as tokens of their victory'" (p. 497). The authors emphasized that this statement was contrary to fact because no promise of any prizes had been initially made.

Kruglanski, et al., (1972) reported that subjects in the prize group rated their task as less enjoyable than their counterparts in the no-prize group ($p < .01$). This tendency persisted into a second week, although it weakened considerably, contrary to expectations. In

addition, on a multiple choice questionnaire, 7 out of 36 subjects in the prize condition group attributed their participation in the games to the prize. As expected, no students in the no-prize group mentioned a prize as a goal for participation. The data were interpreted as supporting a theory of misattribution. Although the prize was introduced only after the games, its eventual offer led students to misattribute the cause of their participation, as well as to significantly lower their task enjoyment. The authors concluded by cautioning against possible detrimental effects of reinforcement by extrinsic factors such as monetary promises, status symbols, or grades, maintaining that rewards serve to lower task enjoyment.

Kruglanski et al. (1972) maintained that an extrinsic award would have a negative effect on a subject's attitude toward a particular task; their definition of extrinsic reward was broadened to include any extrinsic reward presented after the performance of a task, regardless of whether it had been promised or not. This conclusion contrasts to that drawn by Lepper, Greene, and Nisbett (1973) who predicted that only expected rewards would have a negative effect on task behavior.

Kruglanski et al. (1971, 1972) focused attention on some of the contradictory predictions about effects of

extrinsic rewards on motivation. Kruglanski et al. (1971) predicted that extrinsic rewards will have a negative effect on all behavior, not just on behavior that is intrinsically motivated (Deci, 1971, 1972a; Lepper, Greene & Nisbett, 1973). Kruglanski et al. (1972) predicted that even unexpected rewards may have an adverse effect on task behavior, whereas Lepper et al. (1973) have drawn different conclusions for the effects of expected and unexpected rewards. These crucial discrepancies in theory, remain unresolved experimentally to date.

Lepper, Greene, and Nisbett (1973) conducted a field experiment within an educational setting to test what they called an "overjustification" hypothesis. They maintained that a person's intrinsic interest in an activity may be undermined by inducing him to engage in that activity as an explicit means to some extrinsic goal. When the external justification provided is psychologically oversufficient the person will decide that his actions were motivated by external contingencies rather than by any intrinsic interest in the activity itself, thereby decreasing subsequent intrinsic motivation.

Lepper, Greene, and Nisbett (1973) adopted a position that differed from other researchers in this area. Unlike Deci (1971) they maintained that the nature

of the extrinsic goal is not important, and that an overjustification effect can result "even when the reward is insubstantial or merely symbolic" (p. 130). In contrast to Deci (1971) who was concerned with monetary rewards and Kruglanski, et al. (1971) who were concerned with tangible incentives, Lepper, Greene, and Nisbett (1973) focused their attention on the negative effects of symbolic rewards on intrinsic motivation. Although Deci (1971) differentiated between the effects of verbal and tangible reinforcement, Lepper et al. (1973) made no reference to the effects of verbal reinforcement, and it is unclear whether it was to be considered as a symbolic form of reinforcement.

The overjustification hypothesis also differs from the attributional hypothesis postulated by Kruglanski et al. (1972) concerning the effects of unexpected rewards. Lepper, et al. (1973) hypothesized that engaging in an activity for a reward undermines interest in that activity, and that the unexpected receipt of a reward after engaging in an activity should have "little or no detrimental effect" on intrinsic motivation (p. 130).

Lepper et al. (1973) conducted a study in a nursery school. They allowed children to choose play materials from among a large variety that were regularly available and other materials which were only periodically available. For purposes of this experiment, a new and

"novel" activity was introduced, but was made available only on a periodic basis. Children were given the opportunity to draw freely with magic markers. The use of one-way mirrors allowed the experimenters to obtain an unobtrusive measure of the childrens' intrinsic interest in this activity. The index of interest was defined as the percentage of time that the subject chose to play with the experimental activity.

Baseline data were collected during the first hour of three consecutive class days. All children whose total playing time exceeded four minutes of play with the target activity (out of a possible 180 minutes) were selected for study. These youngsters were categorized according to class, sex, and rank in playing time, and were randomly assigned to three treatment conditions.

Children were individually brought into a separate room and asked to draw some pictures. In an expected-award condition, children were shown a sample "Good Player Award" that could be obtained by drawing some pictures for the experimenter. In an unexpected award condition, youngsters were not told about or shown the award until after they had completed their drawings. In a no-award condition, the subjects received no mention of awards. A second experimenter conducted each drawing session, and acted, "friendly, but not overly responsive to the subject" (p. 133). He was blind to the subjects'

condition until the end of the session, when he administered appropriate awards.

Post-experimental observations began seven to fourteen days after treatment. The observational setting and data collection procedures were identical to those used during the baseline period, with observers blind to the experimental condition of the subjects. Due to the dropouts, the final sample consisted of fifty-one subjects.

The results supported the overjustification hypothesis. A one-way analysis of variance showed a significant effect of experimental treatments ($p < .05$). Subjects in the expected-award condition spent a significantly smaller percentage of their time playing with the target materials than students in either the unexpected award group or the no-award group. Subjects in the expected-award condition also showed a significant decrease in interest in the target materials from baseline to post-experimental sessions ($p < .02$), whereas, subjects in the other two groups showed no significant changes in interest levels.

Lepper et al. (1973) interpreted their results as supporting an overjustification hypothesis and contended that their results imposed limitations on the utility of behavior modification techniques. Specifically, they concluded that the institution of token economy schedules

or programs where grades, gold stars or special privileges are awarded to an entire class of children, can be detrimental to some of the students. Since many children will have an initial intrinsic interest in at least some of the activities they are being asked to participate in, the introduction of intrinsic awards will serve to undermine their interest.

Lepper et al. (1973) therefore, argued that the use of tangible rewards can be detrimental to those children who have an initial intrinsic interest in a particular activity. They were careful to point out that they were not unilaterally opposed to the use of token economies. They acknowledged that extrinsic incentives are often used effectively in certain activities. But they maintained that such programs should only be introduced when the level of intrinsic interest is very low.

Lepper, Greene and Nisbett (1973) demonstrated that the expectation of extrinsic rewards may, under some conditions, decrease intrinsic motivation. However, their conclusions about limiting the use of behavior modification procedures are undercut by limitations in their study. The experimental paradigm that they utilized differed in many respects from behavior modification programs. For example, Lepper et al. (1973) utilized a one-trial paradigm, whereas behavior modification programs utilize a multi-trial paradigm.

Lepper et al. (1973) referred to rewards but did not establish whether they were in fact reinforcing. Lepper et al. (1973) described their reward as "non-contingent", but they were presented as contingent upon task engagement. According to terminology defined in this paper, the rewards were "task-contingent". Behavior modification programs typically utilize a performance contingent reward, based on objective standards. No standards of performance were presented or implied in the Lepper paradigm. Due to these differences, efforts to relate the Lepper et al. (1973) results to behavior modification techniques, can be criticized. The crucial and differentiating conditions that produce negative reward effects remained unclear.

Lepper, Greene, and Nisbett (1973) contended that one such condition is the initial level of interest of the subjects. However, their study did not demonstrate that initial student interest was actually the significant differentiating variable. They attempted to select subjects on the basis of their "great initial interest in the drawing activity" (p. 135). Yet, subjects were included in the experiment despite the fact that their playing time was as low as four out of 180 minutes. This comprises less than three percent of the total possible time. Therefore, assertions that high initial interest was evident can be challenged.

It would seem logical to assume, on the basis of the overjustification hypothesis, that the expectation of award would have a greater detrimental effect on those subjects highest in initial level of interest, than on those at the lower end of the continuum. However, this experiment was not designed to test any substantiation of this effect. Whereas the high group declined more than the low group (although no test for statistical significance was reported), the authors noted that this may have been due to a floor effect since the group was initially low in interest. The authors are somewhat aware of this shortcoming of their data. They noted that the range of initial interest was "fairly large" and recognized that,

> It would be of some theoretical interest to know whether the expected award treatment had a different effect on children high in initial interest than on children low in initial interest. Unfortunately the data do not allow a clear answer to this question. (p. 135)

In 1974 Greene and Lepper replicated their original study and in addition studied the variable of contingency. Reward contingency was manipulated by informing the subjects that awards would be given to all who attempted the task (a low performance demand) or to only those children who drew the very best pictures (a high performance demand). In actuality, both groups received the reward.

The experimental materials and methodology were similar to those of the original study: a) Preschool children served as subjects, b) drawing was selected as the target activity, c) a good player award was given and d) classroom measures of time spent on the target task served as the dependent measure. Methodological controls for experimental bias and observer reliability were incorporated into the design. However, no baseline measure of intrinsic interest was obtained.

The contingency manipulation did not produce statistically significant effects, but the original findings were replicated. Subjects in the expected award condition showed less subsequent interest than subjects in the unexpected-award condition. And children in the no-reward control group did not differ significantly from the unexpected reward group.

It was also found that differences in children's performance during the experimental session paralleled differences in their subsequent intrinsic interest. Expectation of an award led them to draw more pictures of a lesser quality (during the experimental sessions), but there were no differences in quality or quantity of pictures drawn related to the contingency manipulation demand for the "best pictures."

The authors concluded that this study lent further support to the notion that,

> The mere presentation of the drawing activity as a means to a salient ulterior goal can be sufficient to produce a decrease in later intrinsic interest in that activity. (p. 1144)

The 1974 Greene and Lepper study replicated the Lepper et al. (1973) experiment, and thus provided further support for the overjustification hypothesis. However, conclusions about the relative effects of task-contingent and performance-contingent rewards can be considered as only tentative because the contingency manipulation was vague. The experimenter did not specify the level to be achieved. The rewards did not produce any significant behavioral changes, and it is possible that the preschool subjects did not understand and/or were not aware of the significance of the instructions.

Lepper and Greene (1975) conducted another replication and extension of the Lepper, Greene, and Nisbett (1973) experiment. This study varied expectation of reward and surveillance. The reward was an opportunity to play with some attractive toys, and was given to both groups. An expected-award group was informed of and shown the toys before the experimental session, but the unexpected reward group was not. In addition, surveillance by a television camera was introduced as a second factor in the design. Subjects in the surveillance group were told that the experimenter

would be monitoring their activity for some part of the session via the camera.

In addition to changing the nature of the reward, from their earlier study, Lepper and Greene (1975) chose a puzzle solving task instead of drawing with magic markers. Rewards were promised only if the child "did a good job on the puzzles" (p. 482). The measure of intrinsic motivation in this study was the proportion of subjects who showed an interest in the target activity rather than the proportion of time spent by each subject on the task. In other respects the paradigm was similar to Lepper et al. (1973) study.

The data revealed that children's expectancy of reward and surveillance yielded significant effects. The authors concluded that this study replicated the findings of previous research showing that intrinsic incentives may undermine childrens' intrinsic interest in an activity. As in their previous research, Lepper and Greene (1975) did not test assumption that a high level of intrinsic interest was crucial.

Greene, Sternberg, and Lepper (1976) designed an experimental research study to directly test an overjustification hypothesis with token reinforcement programs. These authors were cognizant of the important differences between overjustification studies and token reinforcement programs.

The study was conducted in an elementary school with an ongoing individualized mathematics program, that included a biweekly "Awards Assembly". Subjects were 44 fourth and fifth graders. Materials presented to the subjects included four folders, one for each of four different mathematical activities. Time on task was recorded individually by the students, under close supervision.

The first week was devoted to introducing the four new activities, one each day, with the fifth day devoted to the administration of a questionnaire about preferences of the four different activities. The baseline phase was conducted during the next 13 school days, and students played with the four activities without differential reinforcement.

Subjects were then assigned into one of four groups on the basis of the time they had concentrated on their most preferred activities. The four groups were high interest, low interest, choice group, and control group. The "high interest" group was rewarded for time spent on the two activities with which they had spent the most time during baseline. The "low interest" group was rewarded for time spent on the two activities with which they spent the least time. Subjects in the "choice group" were allowed to choose two target activities. The control group was rewarded for time on task for any of

the four activities. Procedures were designed in a manner such that none of the personnel involved in the study were aware of students' experimental condition. The treatment condition continued for 13 consecutive days.

The reward consisted of credit on students' pre-existing math award sheets; one credit point for every three hours of work on their target tasks. Credits were redeemable for certificates and trophies presented at the Awards Assembly. The dependent measure was the amount of time spent playing with the target activities.

For the withdrawal phase, students were told that other youngsters in the school resented the unfair advantage they had, and all reinforcement contingencies were withdrawn. The math lab materials were left as part of the curriculum, and students were encouraged to continue using them and/or to make up new materials.

It was found that each of the three experimental groups were reinforced by the credits. Those differences were confined to the time-on-task measure. No systematic differences in accuracy, rate of performance on target tasks, progress in the school's regular math program, or on subject's responses to a final questionnaire, were found. During the withdrawal phase, subjects in each of the three experimental groups spent less time with their target activities than they had during the baseline

phase. This post-treatment drop below baseline was statistically significant for the "high interest" and "choice group". It did not reach statistical significance in the "low interest groups" possibly because of the restricted range between initial baseline and zero. Between group comparisons indicated that during the withdrawal phase, children in the "low interest group" and in the "choice group" spent significantly less time with their target activities than did control subjects matched or yoked to them. However,there was no significant difference between the "high interest" and "control group" during withdrawal.

The authors concluded that under some conditions, multi-trial contingent reinforcement procedures are capable of producing post-treatment decrements in task engagement. They concluded that there was "clear evidence that overjustification effects are not limited to single-trial, 'non-contingent' reward procedures" (p. 1229). They acknowledged that their results did not permit unambiguous conclusions to be drawn concerning the effects of differences in subjects' levels of initial interest. Based on previous research they had expected overjustification effects to be more likely among subjects rewarded for engagement in activities of relatively higher initial interest, yet this effect did not occur.

Greene, Sternberg, and Lepper (1976) demonstrated that there can be negative results of extrinsic rewards, even within a multi-trial token economy program. And there were indications that negative effects are not necessarily limited to tasks of initially high intrinsic interest. It should be noted that in the Greene et al. (1976) study, rewards were presented contingent only upon time on task. They were not contingent on any factor which might reflect increased competency, such as number of problems correctly solved. The reward is designed as a task-contingent reward. The differential effects of a task contingent and performance contingent reward may be a delineating factor.

The studies of Lepper and colleagues (1973, 1974, 1975, 1976) have demonstrated that the expectation of a reward, may in some instances, have a negative effect on the subsequent performance of an activity. It remains unclear, however, what specific conditions bring about this effect. Although overjustification theory assumes that the negative effect of reward is directly associated with initial high intrinsic interest in the target activity, this assumption has not been experimentally verified.

Calder and Staw (1975) designed the first reported study to test the hypothesis of the additivity versus the interaction of intrinsic and extrinsic motivation.

Calder and Staw (1975) experimentally manipulated intrinsic as well as extrinsic factors, and the effects and interaction of these dependent variables on relevant independent variables was measured. They predicted an interaction between intrinsic motivation and extrinsic rewards. Specifically, they hypothesized,

> ... an inverse (or self-perception) effect when a task is initially high in intrinsic interest and a direct (or reinforcement) effect when there is initially less interest in a task. (p. 600)

Calder and Staw (1975) acknowledged that the concept of intrinsic motivation was unclear as a psychological construct but maintained that it was useful at a descriptive level and that "it may be viewed as a perception on the part of individuals" (p. 599). Extensive pilot testing was conducted in order to find two parallel tasks that could be validly labeled as pleasurable and unpleasurable.

The basic experimental materials finally selected consisted of 15 jigsaw type puzzles, with the manipulation of intrinsic motivation accomplished by having the puzzles blank for one group of subjects, versus having interesting pictures on the puzzles of another group. Differences in possible overt behavior of the college subjects was controlled by making both sets of puzzles "extremely routine" with "no possibility of making an error" (p. 601). Pilot testing of attitudes

towards the puzzles seemed to confirm that there were no cognitive differences, and that only affective differences, such as interest and pleasure, were significant. The authors acknowledged that it was impossible to prove that there are no unintended differences in overt behavior or perceptions, but maintained that their prototype design minimized this possibility.

Extrinsic motivation was manipulated by rewarding the experimental group with a dollar, whereas, the control group was neither promised nor given any reward. The authors conceived of the reward as "equitable, salient, expected, non-contingent, and given at the end of the task" (p. 602). The dependent variable of task satisfaction was measured by an attitude scale that was presented after the completion of the assigned task. A behavioral measure of motivation was not obtained since the procedure did not lend itself to free time assessment. The closest they were able to approximate a behavioral measure of motivation involved recording students' willingness to volunteer for future experiments of a similar nature, without pay.

The enjoyability ratings were analyzed and a significant task by reward interaction was found. The interaction was disordinal, such that with the introduction of money interesting puzzles were rated less

enjoyable than boring puzzles; the authors note that this is stronger than theoretically required to support their hypothesis. The behavioral questionnaire revealed no statistically significant interaction, however the pattern did parallel those of the enjoyability ratings.

This study was the first to test for and find a statistically significant interaction between the effects of reward and initial task interest. Although the sample was relatively small (total N = 40), it was well controlled with stringent care taken in the design of the independent variables. It should be noted, however, that the dependent measures were both attitudinal, with no behavioral counterpart.

The experimental research studies by Deci (1971, 1972a, 1972b), Kruglanski et al. (1971, 1972), Lepper et al. (1973, 1975), Greene et al. (1974, 1976), and Calder and Staw (1975) repeatedly demonstrated that extrinsic rewards could have a negative effect on subsequent intrinsic interest. They were unable, however, to provide consistent information as to the specific precedent conditions that elicited these negative consequences. The focused attention on the possible negative effect of rewards and stimulated further research on the effects of extrinsic rewards on subsequent motivational levels.

Experimentally Challenging Research

Advocates of behavior modification techniques (Feingold & Mahoney, 1975) challenged the generalizability of the overjustification hypothesis. They contended that token economy programs do not necessarily result in a loss of intrinsic interest, even if initial interest levels are high. They sought to support behavior theories with experimental research.

Feingold and Mahoney (1975), in an experimental study entitled, Reinforcement effects on intrinsic interest: Undermining the overjustification hypothesis, effectively did just that. They charged that the procedures used by Deci (1971, 1972a) and Lepper et al. (1973, 1975) were not parallel to those typically encountered in classroom token economies, and strongly doubted the external validity of these studies. They designed their experiment with an emphasis on three methodological considerations:

> (1) actual demonstration of a reinforcement effect (performance increment) relative to baseline measurements, (2) close parallels to the actual procedures employed in many classroom token economies, and (3) sufficient continuing measurements to examine temporal trends and transition states. (p. 369)

Five second grade children were given access to an etch-a-sketch game, and follow-the-dot picture completion books, and freedom of choice to play with either (but with nothing else). During eight baseline sessions of

fifteen minute duration, the performance of the children in the dot-to-dot books were recorded without their knowledge. A token reinforcement phase followed during which subjects earned points for exceeding their highest baseline performance. Points were exchangeable on a daily basis for toys displayed in the back of the room. Token reinforcement procedures were continued for four sessions.

Baseline 2 immediately followed the reinforcement period. Subjects were told that they would no longer receive prizes, and interest in their work was no longer shown. Baseline 2 was conducted for eight sessions over a two week period. Baseline 3, was a test of delayed results, and was conducted after a lapse of two weeks of no sessions.

Data were subjected to repeated measures analysis of variance with two planned orthogonal comparisons. The first planned comparison verified that a significant increase in level of performance occurred during the reinforcement period ($p < .001$). The second planned comparison of the last sessions of Baseline 1 and the first sessions of Baseline 2 showed no significant differences, indicating no immediate changes due to the effects of reinforcement. (Overall F was significant). A subsequent comparison of performance on Baselines 1 and 3, showed that despite wide variability, there was a

consistent increase in the average performance for each of the five subjects. Contrary to the overjustification hypothesis, the children displayed an average increase of 95.08 percent in performance. (No statistical analysis of this comparison is reported.)

Feingold and Mahoney (1975) concluded that the results of this experiment "cast doubt on the assertion that extrinsic reward necessarily undermines intrinsic motivation" (p. 375). They showed that a repeated session, token economy program similar to that used in many classrooms, produced no subsequent deterioration of intrinsic interest.

It is noted, however, that this study had only five subjects, with each subject acting as their own control. The statistical analysis led to the acceptance of the null hypothesis; it showed only that reinforcement during the experimental session led to no subsequent decrease in overall performance. Regrettably, the most interesting comparison, that of Baseline 1 and 3, was not statistically evaluated. If a post hoc comparison of these two baselines was in fact significant, it would have indicated that reinforcement eventually had a positive rather than negative influence on intrinsic interest.

The Feingold and Mahoney study (1975) challenged the assertion that overjustification effects will occur

whenever the task is apparently intrinsically interesting. However, it remains unclear concerning what variables determine when overjustification effects will or will not occur. There are inherent differences between typical paradigms followed by attribution theorists and the Feingold and Mahoney (1975) token reinforcement paradigm. Unlike attributionist research designs, prior to 1975, Feingold and Mahoney (1975) presented the reward as contingent on attainment of a specified objective level. And unlike any attribution study to date, the standards were established individually for each subject on the basis of baseline levels. The importance of these standard setting conditions, and the extent to which they influenced intrinsic motivation, has yet to be determined.

Reiss and Sushinsky (1975), in a two part experimental series, challenged the generalizability of the overjustification hypothesis (Lepper et al. 1973, 1975). They contended that the Lepper and associates' paradigm is a one trial non-contingent promise of reward situation, and their results would differ from a multi-trial contingent reinforcement condition. The difference is an essential one, since the multi-trial paradigm more closely parallels procedures employed in behavior modification programs.

The first experiment was designed as a replication and extension of the Lepper et al. (1973) paradigm, with the experimental manipulation of promise of reward and exposure to distracting stimuli. The dependent variable was the proportion of time spent on the target task during posttest.

The task chosen in this study differed from that of both the Lepper et al. (1973, 1975) studies. The only restriction Reiss and Sushinsky placed on their choice of an activity was that it be of a playful nature, and therefore be intrinsically interesting. The target task involved listening to a casette recorder for a five minute period until a buzzer sounded. The opportunity to play with an attractive doll was the promised reward. The exposure to the doll prior to listening was conceived of as a distracting stimulus.

During the individual experimental sessions, each of 32 first grade girls was instructed to listen carefully to the song until the buzzer sounded. Depending on the group to which the children were assigned, they were promised rewards and exposed to the doll. After listening to the song for five minutes, all subjects were permitted to play with the doll for three minutes before being returned to the classroom.

Later in the same day, the five minute posttest was conducted. It was a free-play situation in which the

child could choose to listen to any of the three songs. Four children, one from each experimental condition, were brought into the room at the same time; all had been reinforced for listening to the same song. Two observers in the room recorded the length of time each child spent actually listening to the target song. Interrater reliability was near perfect ($r = .99$).

It was found that expected awards produced less listening time to the target song during the posttest sessions. This study replicated the results reported by Lepper et al. (1973).

The second experiment was designed to test for an overjustification effect following termination of a token economy program. The implication is that this paradigm parallels procedures commonly employed in behavior modification programs, the essential difference being only that a form of play is the behavior to be modified.

Nine kindergarten subjects were trained to discriminate three songs assumed to be of equal interest. Three children were trained together, but each was reinforced for listening to a different target song. During each of ten reinforcement trials, the children earned tokens for listening to the target song. There was no reinforcement for listening to the other two songs. The tokens could be traded for a toy at the end of the session.

The posttest was conducted two days later, in a setting similar to that in the first experiment. Three subjects, each trained on a different song, were brought into the room. They were instructed to do whatever they wanted to, and were free to listen to any of the three recordings. The listening time for each song was recorded by two observers.

Results showed a statistically significant preference for the target song designated during experimental training compared to non-target songs. The mean time spent listening to the target song during the 600 second post-test was 350 seconds, compared to a mean of 145.6 seconds spent listening to the most preferred non-target song. Reiss and Sushinsky (1975) concluded that "this finding disconfirms the overjustification hypothesis prediction of decreased interest in the target song, indicating instead an opposite effect" (p. 1122).

Reiss and Sushinsky (1975) have presented a theoretical and experimental challenge to the Lepper et al. (1973, 1975) overjustification hypothesis. This series of experiments were attempts to both replicate the Lepper et al. (1973) experiment, and to differentiate that paradigm from those used in behavior modification studies. Experiment I was designed as a replication of the Lepper et al. (1973) paradigm, and experiment II as a replication of the behavior modification paradigm. Reiss

and Sushinsky (1975) isolated the number of trials of training as the essential differentiating characteristic. Theoretically, the only differentiating factor between experiment I and II should have been in the number of trials of discrimination training. However, in actuality there are many differences in methodology between the two experiments, and consequently, it is not possible to conclude that the contrary results are due primarily to the number of trials of training.

There were many differences between the two experiments. In experiment I the discrimination training was conducted individually whereas in the second experiment the children were trained in groups of three,each of whom was reinforced for listening to a different song. Posttesting was conducted in small groups in both experiments, but in experiment I children who had been reinforced for listening to the same song were tested together, whereas in experiment II children who had been reinforced for listening to different songs were tested together. Posttest delay also ranged from a three to five hours in the first study as opposed to a 48 hour delay in the second study. Any of these differences may be the crucial factor, and difference in results between the two experiments cannot simply be attributed to the difference in the number of trials of training.

For example, it remains questionable whether the methodology in experiment II allowed for a test of the overjustification hypothesis. This theory is based on the notion of cognitive re-evaluation, i.e., "if I am rewarded for doing this it must not be interesting". However, in the paradigm employed in the second and crucial study, three children received training in the same room, and each was rewarded for listening to a different song. Therefore, it does not seem logical even according to the overjustification hypothesis, for the subjects to conclude that one song is of less intrinsic interest than another. The change in methodology seriously confounded interpretation of results.

In addition to these basic changes in the methodology, the second experiment is flawed by the lack of a control group. Without a control group the results demonstrate only that the token reinforcement program results in an increase in listening to a target song over nontarget songs, but not that the increase was due to the token reinforcement. The increased interest may have been a result of increased familiarity, rather than due to the token program of reinforcement. It may be that familiarity with a particular song leads to increased to increased interest in that song, in kindergarten children, and that interest in the target song would have been even higher had tokens never been introduced. A

control group is an essential part of any scientific experiment, and due to the differences in methodology as well as in sample, the subject population of the first experiment can in no way be considered as a replacement for a randomly selected control group.

The differences in the sample population of the two experiments is a factor which is difficult to understand. Even though the second experiment is the more crucial one, only nine subjects were used, whereas the first experiment had 32 subjects. Moreover, the second experiment used both male and female students, while the first experiment used only girls. No explanation is given for these differences in subject choice.

The task chosen, listening to a song, is again assumed to be intrinsically interesting. No verification of this assumption is attempted and in this instance the assumption of intrinsic interest is not logically supported. It is difficult to believe that all children would find listening to a song for a five minute period intrinsically interesting. Five minutes is a long time for a kindergarten child to remain inactive, and often exceeds the attention span of even an older child. If this is indeed so, then the interpretation of results is even more difficult. Consequently, it is concluded that Reiss and Sushinsky (1975) have neither confirmed nor disconfirmed the overjustification hypothesis. Their

challenge remains theoretically defendable but experimentally unsubstantiated.

Vasta, Andrews, McLaughlin, Stripe, and Comfort (1979) designed a study to examine possible undermining effects of reinforcement when procedures employed more closely resembled actual token programs within the classroom. They incorporated four characteristics into their design: children were studied in groups, the experimental setting involved a "teacher" and a "class", the children were free to move about the room, and reward involved both social and tangible reinforcers. In addition, two groups of subjects were studied, one group displaying initial low level behavior in the target task, and the other group a relatively high level of initial interest.

Participants were six children of approximately seven years of age, who were escorted by a "new teacher" to a small classroom. The children were told that they could engage in any of three activities in the classroom: blocks, cardboard puzzles, or coloring book pages. Each session lasted 15 minutes. Data were collected by observers working behind two way mirrors, (observer reliability was 94.3%).

During baseline sessions, the experimenter circulated around the room without initiating contact with the children. During reinforcement sessions the

experimenter observed each child for a twenty second interval, following a list of randomly organized names. If at the moment of observation the child was engaged in the target task, he/she was reinforced. The target task was coloring; the reinforcement was a star drawn on the child's name tag and verbal praise. Posttreatment consisted of a return to baseline procedures. Follow-up conducted after a four week interval again involved only baseline procedures.

The results indicated that time on target activity rose quickly with the introduction of the reinforcement and remained high throughout the Posttreatment and Follow-up phase. A repeated measures analysis of variance reached significance for these data ($p < .01$). There were no differences in results obtained between a first and second experiment, although six different subjects were used, and baseline performance on the target task differed between the groups (m = 30.8% to m = 63.5%).

The authors concluded that prohibitions against the use of reinforcers within classroom token economy programs may be unwarranted. They noted that even within their second experiment, where initial interest in the target task was high, there was no inevitable drop in nonreinforced response rate. Nevertheless, they recommended caution be exercised in the use of

reinforcement programs, until more empirical evidence was available.

Vasta et al. (1979) noted that they did not specifically design their experimental research to address the overjustification hypothesis. Consequently, it remains difficult to interpret their results as either supporting or refuting such a formulation. Their study is confounded by the utilization of both symbolic, and undefined verbal praise. And their findings of no differences between groups initially low and high in intrinsic interest on the target task does not clarify any theoretical issues. Vasta et al. (1979) have demonstrated that reinforcement does not inevitably lead to a loss of intrinsic interest. Unfortunately, they have not delineated the critical variables that determine when reinforcement will be beneficial or detrimental. The experimental findings of Feingold and Mahoney (1975), Reiss and Sushinsky (1975), and Vasta and his colleagues (1979), demonstrated that within multi-trial paradigms, and extrinsic reward did not result in a loss of intrinsic interest, even when initial interest levels were high. However, none of the above were able to delineate the conditions which differentiated between the negative and positive effects of extrinsic reward. None were able to demonstrate that the critical variable was multi-trial versus single-trial conditions. None was

able to demonstrate that the critical variable was high versus low initial interest in the target task.

Reward Contingency and Competency

Deci (1975) hypothesized that rewards conveying information about competency should sustain rather than undermine intrinsic interest. Experimental investigations followed which were designed to demonstrate that when rewards were presented in a manner such that they conveyed a message of competency there would be no subsequent loss of intrinsic interest.

Karniol and Ross (1977) contended that discounting of intrinsic interest may not be an inevitable result of reward dispensation. They designed their study to delineate some of the conditions under which rewards may not be detrimental to childrens' intrinsic motivation, and they directly examined the relation between reward contingency and intrinsic interest. They maintained that:

> Extrinsic factors which define one's performance as competent should generate intrinsic interest in the activity. Thus, when rewards provide the standard by which the individual evaluates the quality of his performance, intrinsic interest should be generated rather than discounted. (p. 482)

Their study was based on the formulation of a number of theorists (deCharms, 1968; Smith, 1968; White, 1959) who have argued that a feeling of competence is the precursor of intrinsic interest. Organisms are perceived as

motivated and as deriving intrinsic satisfactions from those activities at which they feel competent. And they may judge their competence by extrinsically and socially defined standards.

Previously published research depicting the negative effects of rewards on subsequent intrinsic interest focused mainly on task-contingent rewards. Those few studies that included performance-contingent rewards found them to be equally detrimental. Deci (1971) found that performance-contingent rewards did decrease intrinsic interest. Karniol and Ross (1977) countered that in these studies it was "not clear whether subjects perceived themselves as competent or effective at the task because no standard of performance was established" (p. 483). Greene and Lepper (1974) included a manipulation of performance demands in that some subjects were promised a reward only for good pictures. These subjects also showed a subsequent loss in intrinsic interest. Karniol and Ross (1977) interpreted these negative findings as inconclusive, since "Greene and Lepper failed to obtain any evidence that the experiment successfully manipulated performance demands" (p. 483). Karniol and Ross emphasized that:

> Performance-contingent rewards should maintain intrinsic interest only to the extent that they clearly convey information about one's effective performance. (p. 483)

The Karniol and Ross (1977) research included an experimental manipulation of reward, as well as a manipulation of the degree to which subjects succeeded at the activity. Reward contingency was manipulated by assigning subjects to either a "performance-relevant reward" condition, a "performance-irrelevant reward" condition, or a "no-reward" (control) condition. The "performance-relevant reward" is described as "a reward contingent on how well they did at the task" (p. 483). For the sake of consistency, it will heretofore be referred to as a "performance-contingent" reward. A "performance-irrelevant reward" is defined as "disbursed for task engagement per se" (p. 483). It will be referred to as a "task-contingent" reward. To manipulate feelings of success or failure, subjects in the "low performance" condition were told that the average child of their age chose 16 correct answers, and that a performance of less than 16 was below average. Subjects in the "high performance" condition were told that if they got more than six correct they were better than average, but if they got less than six correct they were below average.

Manipulation of the reward contingency was planned so that all rewarded subjects received two marshmallows. Subjects in the "task-contingent" group were simply told that they would receive two marshmallows. Subjects in

the "performance contingent" group were told that the number of marshmallows they would receive would depend on their performance. In the "high-performance" condition they were promised one marshmallow for less than six correct responses, and two for six or more correct responses. In the "low-performance" condition subjects were promised three marshmallows for 16 or more correct responses, but only two marshmallows for less than 16 correct responses. Subjects in the control group were neither promised nor did they receive any marshmallows.

Subjects were 57 children, ages four to nine, who attended a summer playground program. Subjects were individually brought into a trailer to play a slide game. In actuality, the game had no correct or incorrect answers, but a tape programmer provided subjects with feedback information such that all subjects had 10 out of 20 correct responses.

Intrinsic motivation was measured during a free play period, immediately following the experimental session. Subjects were allowed to choose between the slide game and four alternate toys, and they were informed that no rewards would be given. Free choice time of the slide game during a six minute session served as the dependent measure of intrinsic interest.

As predicted, a significant contingency X level of performance interaction ($p < .005$) was found. Under the

"high performance" condition, the task-contingent reward led to a significant decrease in play with the slide game relative to the performance contingent and the control group ($p < .01$). Under the "low performance" condition, play with the slide game was significantly reduced for the control group, and there was a similar trend for the performance-contingent reward group.

Karniol and Ross (1977) interpreted their results as providing support for the hypothesized relationship between reward contingency and intrinsic interest. Specifically, a performance-contingent reward, when presented under success condition, does not lead to a decrease in intrinsic interest. The results indicated that when the child has succeeded only a task contingent reward decreases subsequent intrinsic interest.

> The results suggest that when the child has succeeded a performance-irrelevant reward undermines interest because it de-emphasizes the importance or significance of one's success and the intrinsic satisfactions that accrue thereby. (p. 485)

Karniol and Ross (1977) designed their research to compare directly the results of a task-contingent and a performance-contingent reward. They maintained that a performance-contingent reward can convey a message of competency, and the consequently no loss of intrinsic interest will follow. They further stipulate that the reward can convey competency only when it conveys

information about ones degree of success. Therefore, the performance-contingent condition utilized by Karniol and Ross (1977) specifically differed from the performance-contingent condition in the Greene and Lepper (1974) study, in that standards were clearly defined for all groups prior to task engagement. It should be noted that their standards simultaneously conveyed within subject and between subject standards, i.e., they were given an objective definition of success, as well as a way of comparing their level of success to that of their peers.

The Karniol and Ross (1977) research lends empirical support to the contention that a tangible reward can be contingently presented in a manner such that there is no loss in subsequent intrinsic interest. The author's interpretation of these phenomena is theoretically interesting and merits further experimental validation. Direct measures of those feelings attributed to subjects, would be most informative. The negative effects of a reward was ascribed to the ability of a task-contingent reward to undermine interest by de-emphasizing the importance of the persons success. Karniol and Ross (1977) did not measure their subjects' perceived feelings of task importance. The positive effects of a reward was viewed as related to the ability of a reward to convey a message of competency, and the importance of this

intervening variable was repeatedly stressed. Nevertheless, Karniol and Ross (1977) made no attempt to directly measure the theorized changes in levels of perceived competency.

Enzle and Ross (1978) presented and experimentally tested their interpretation of Deci's (1975) hypothesis regarding the differential effects of reward. Deci (1975) maintained that a reward can have both control and informational attributes. He predicted that when the control aspect is salient, a reward will decrease intrinsic interest. However, when the competence information is salient, a reward will increase peoples intrinsic interest in the task. Enzle and Ross (1978) contended that the control aspect is salient when the reward is offered contingent solely on the execution of an activity (task-contingent reward). The informational aspect of a reward becomes salient to an individual when he believes that the receipt of the reward is contingent on successful performance of the task (performance-contingent reward).

In the Enzle and Ross (1978) study, subjects were monetarily rewarded on a simple task (task-contingent), for achieving a criterion performance level (performance-contingent), or non-contingently (a control condition). The researchers also investigated the effects of different levels of reward level on intrinsic

interest, hypothesizing that a reward of trivial magnitude "should neither instigate the reattribution process nor enhance feelings of competence" (p. 591).

Enzle and Ross (1978) used a 2 X 3 factorial design with the experimental manipulation of payment level, high versus low, and reward contingency, unexpected versus task-contingent versus criterion-contingent (hereafter referred to as performance-contingent). Subjects, 78 college men, were randomly assigned to condition. The target task was a four piece non-commercial wooden block puzzle, with cards depicting three configurations which could be constructed from the puzzle pieces. Pilot testing indicated that all subjects could complete the puzzles in not less than three minutes and not more than twelve minutes.

The study was presented to students as creative problem solving. It was individually administered. Subjects were instructed that they would have eight minutes to complete each design. In the task-contingent condition,

> The experimenter clearly explained that receipt of the payment depended solely on working at the task, not on quantity or quality of performance. (p. 592)

In the performance-contingent condition, subjects were instructed that they would receive payment,

> ... if their performance reached a skill-related criterion level. This criterion level, although unspecified, had purportedly been determined in a

> pilot study with another group of subjects and reflected a standard of performance and ability well above average. Performance was to be judged by progress made in problem solving, not necessarily by the completion of the tasks. (p. 592)

Money was not mentioned to the noncontingent reward (control) group. The amount of payment was $.45 in the low payment condition, and $1.50 in the high payment condition.

Upon completion of the task all subjects were rewarded monetarily. The amount of payment was varied experimentally. Students in the task-contingent reward group were simply given their appropriate payment. The performance-contingent reward group "received feedback sheets indicating that their performance had reached the criterion level" (p. 592), and were then given the appropriate payment. The unexpected reward (control) group, were

> ... given a memo informing them that surplus funds in the department were being distributed among students participating on any one of several experiments, including the experiment in which they were taking part. The memo clearly indicated that the money was not associated with their performance on the puzzle task. (p. 592)

They were then given appropriate payment.

The dependent measure of intrinsic interest was the summed scores of two questionnaire items. The questionnaire, completed by subjects immediately after receiving payment, included also several questions about

problem solving strategies, designed as a cover story, as well as two measures of intrinsic interest in the target task.

The results indicated a significant interaction between payment level and reward contingency. Under task-contingent reward conditions, high pay led to a loss of intrinsic interest compared to low pay. Under performance-contingent reward conditions, the opposite occurred: high pay led to greater intrinsic interest than low pay. These differences were statistically significant. In the unexpected reward condition, there were no differences between men receiving high and low payments. In the low payment condition, there were no statistically significant differences between reward contingency conditions.

Enzle and Ross (1978) interpreted their results as supportive of Deci's (1975) cognitive evaluation theory:

> Evidence was obtained that contingent rewards may either decrease or increase intrinsic interest, depending on whether the control aspect or the competence information aspect of the reward is made salient... Intrinsic interest decreased when the control aspect of the reward was emphasized and increased when positive competence information was conveyed by the reward. (p. 594)

Results of the Enzle and Ross (1978) study must be interpreted cautiously, however, since the dependent measure was attitudinal. There was no behavioral measure of motivation. Interpretation of results must also take

into account the nature of the contingency manipulation, which differed from other studies. First, the control group was an unexpected reward group, rather than an unrewarded group. Second, the performance-contingent reward was based on a standard of performance which was allegedly normative, but which remained undefined. Third, students in the performance-contingent reward group were unique in receiving direct information concerning the success of their performance. The extent to which this positive feedback may have been the crucial variable was unknown.

In summary, although Enzle and Ross (1978) demonstrated that a trivial reward was meaningless in relation to subsequent intrinsic interest in the target task, they did not offer criteria for distinguishing between a trivial and meaningful reward. They demonstrated that performance contingent rewards could increase intrinsic interest and that task contingent rewards could decrease intrinsic interest. Like Karniol and Ross (1977) they assumed that this outcome was based on the informational and competence enhancing aspect of a performance contingent reward. However, they have made no attempt to measure competency feelings directly. Competency was an unmeasured intervening variable.

Boggiano and Ruble (1979) sought to test the competency hypothesis by manipulating the availability of

direct information about success. They noted that previous support for the competency hypothesis was evidenced on both behavioral (Karniol & Ross, 1977) and attitudinal (Enzle & Ross, 1978) measures, where a reward made contingent on meeting a performance standard produced more subsequent interest than a reward made contingent on task engagement. They maintained that:

> Although intrinsic interest should be maintained when reward attainment is made contingent on an absolute performance standard, the competency hypothesis would predict that if more direct information about competence or incompetence is additionally provided (in this case, performance level relative to others), subsequent level of interest should vary directly with the type of information regarding competence that is presented. (p. 1463)

In more generalized terms, the competency hypothesis predicts that information regarding success should sustain rather than undermine subsequent intrinsic interest.

The competency beliefs were also studied developmentally. According to Veroff (1969), children do not derive feelings of competence from comparative standards of excellence until the age of seven or eight years. However, even young children, such as preschoolers, are assumed to make evaluative judgements regarding their level of competence based on absolute performance standards. Boggiano and Ruble (1979) therefore predicted that competency information conveyed

in relative terms would mitigate the undermining effect of reward for older children but not for younger children.

The major hypotheses of the study were examined in a 2 X 2 X 3 factorial design, with one additional cell. The assigned variable was grade (nursery v. elementary school children); the experimentally manipulated variables were reward contingency (performance-contingent reward vs. task-contingent reward) and social comparison (relative competence vs. relative incompetence vs. no-information-regarding-competence). As a baseline control group, a no-information-regarding-competence, no reward group was included.

Subjects were 147 children of either nursery school age or elementary school age. The target task, "find the hidden pictures", was found to be "highly interesting for the two age groups" (p. 1464). The high interest was indicated by the high percentage of time that the subjects in the baseline control condition played with the task. To make performance level comparable across age groups, the actual difficulty of the task was increased with grade level so that all children obtained a score of four (two exceptions were excluded from data analysis). Alternate tasks were marbles, crayons with a game that involved following the lines to complete a

picture, two mazes, and an additional set of hidden pictures. The rewards were two Hershey's candy kisses.

Children were individually escorted to a mobile trailer, where they were introduced to the target task. All subjects were then told that a score of 3 to 8 would indicate that they had done "okay". Stars were painted above the numbers 3 to 8 on a scoreboard, to ensure the understanding of the absolute criteria. Subjects in the performance contingent reward group were told that they could earn the candy, "only if they obtained a score of at least 3 out of a possible 8" (p. 1464). Children in the task-contingent reward group were told that they would be given the candy "for simply playing the game, regardless of their score" (p. 1464).

To manipulate social comparison information, the scoreboard was present in all conditions. Children in the relative incompetency condition were shown a scoreboard indicating that most of the other children had obtained a score of 7 out of the maximum 8. Children in the relative competency condition were shown a scoreboard indicating that most of the others had obtained a score of 1 out of a possible 8. Children in the control condition were given no comparative scores; they were simply told to enter their score on the board.

After receiving appropriate instruction, children were given the choice of whether or not to play the game.

After completing the game, all children except those in the control group were given their reward.

A behavioral measure of free play was taken immediately afterward. Children were told that they would not be given any further candy during free play. They were studied through a one way mirror for six minutes.

Results indicated a main effect for contingency, with intrinsic interest higher for children in a performance-contingent group than those in a task-contingent group. This effect was primarily due to the data of the younger children. A significant Grade X Social comparison interaction showed that relative competency information affected the subsequent interest level of the older, but not the younger children. For the younger children, the task-contingent reward undermined interest; the task-contingent reward produced less interest than the performance-contingent reward, and the comparative competency information had no effect on their later interest in the task. For older children, intrinsic interest differed as a function of the comparative competency information, in the predicted direction. The effects of reward contingency manipulation (task-contingency versus performance-contingency) had no effect for the older children when comparative information was given. When no comparative

information was given task contingency reward undermined intrinsic interest.

Boggiano and Ruble (1979) interpreted their results as supportive of the hypothesized effect of different types of competency information on intrinsic interest. They viewed it as a replication of the Karniol and ross (1977) study indicating that reward based on successfully meeting an absolute performance standard does not undermine interest. And they maintained that their study provided support for the hypothesis that:

> Information about competence can mitigate the undermining effect of reward on intrinsic interest. The availability of direct information regarding competence presented by means of social comparison appeared to override completely the effect of reward on behavior. (p. 1467-8)

They also interpreted their findings as supporting the hypothesis that young children are not affected by comparative competency information. They suggested that Greene and Lepper's (1974) failure to support the competency hypothesis was due to their preschoolers inability to appreciate social comparison information.

Boggiano and Ruble (1979) assumed that the mediating factor accounting for their findings was the "differential perception of competence resulting from information in the various competency conditions", although they acknowledged that "perceptions of competence were not assessed here" (p. 1467). Boggiano

and Ruble (1979) distinguished between the effect of absolute and comparative standards, demonstrating possible differential effects of absolute and comparative standards according to developmental age. Specifically, they showed that young children can understand and react to absolute standards of competency, but not necessarily to comparative standards of competency.

The Boggiano and Ruble (1979) research study replicated the Karniol and Ross (1977) study. In both experimental paradigms, absolute standards were presented to all subjects, and reward contingency was subsequently manipulated. Under these conditions, only a task-contingent reward undermined subsequent intrinsic interest as measured by time on target task during free play. The extent to which the outcome was affected by the initial presentation of standards remains unknown; the effects of reward contingency manipulation, without any standards specified at the outset, was not demonstrated.

The Boggiano and Ruble (1979) study also served to lend further support to the hypothesized construct of competency feelings as an intervening variable between reward contingency and subsequent intrinsic interest.

Harackiewicz (1979) studied the effects of reward contingency and positive performance feedback on subsequent intrinsic motivation for an enjoyable task,

with the goal of methodologically separating the confounding effects of material reward from positive feedback about performance. It was predicted that positive feedback would enhance intrinsic motivation, that task-contingent rewards would undermine it, and that these two effects would be independent. It was further hypothesized that a performance contingent reward would be more controlling than informational, that it would decrease intrinsic interest more than task-contingent rewards, and that this effect would be greatest when norms were not supplied at the outset.

Subjects were 93 high school students tested in their English class. The Harackiewicz (1979) study was conducted in three phases. The first and third sessions were group administered in the classroom, the second session was individually administered in a separate room. The sessions were approximately one month apart.

The target task was a series of hidden figure puzzles in which the name Nina is hidden in several places. The goal was to locate the Nina inscription. In the first pretest session, subjects were presented with booklets containing Nina puzzles, as well as two other puzzle type tasks. After each puzzle set, subjects rated them for interest and enjoyment. Pretest enjoyment was interpreted as indicating a high level of initial enjoyment of puzzles. The rewards, two felt-tip pens and

a notebook, were delivered at the conclusion of the experiment.

The experimental manipulation and immediate posttest were conducted during the second session. The reward contingency was manipulated three ways: no reward, task-contingent reward, and performance contingent reward. In the no reward condition, none was promised; subsequent to collection of behavioral and attitudinal data, subjects were informed that there would be a gift for "having participated in psychological research" (p. 1357). In the task-contingent reward condition, subjects were promised a reward "for doing the puzzles" (p. 1357). In the performance contingent reward condition, subjects were "promised a reward for exceeding the performance of the average high school student" (p. 1357).

Reward contingency was manipulated relative to positive performance feedback and no feedback. Positive performance feedback was defined as providing standards after the completion of the target task. Positive performance feedback was presented verbally after the completion of three puzzles: "We've found that the average high school student usually finds four Ninas, so you did better than the average student on these puzzles" (p. 1357). The effects of reward contingency and positive performance feedback were examined independently in the task contingent and no reward conditions.

Harackiewicz (1979), however, found "it difficult to separate reward manipulation from feedback in the case of performance-contingent rewards" (p. 1354). Consequently, there was a missing cell (performance-contingent reward with no performance feedback). For the two performance-contingent groups, both received positive performance feedback after the completion of the task, with experimental manipulation differing in whether or not norms were additionally supplied prior to task engagement. Of the six conditions, only subjects in the "performance-contingent reward, norms supplied, positive feedback" group, were informed of the exact performance standard prior to task engagement: "detecting four or more Ninas in the three puzzles combined" (p. 1357).

Attitudinal and behavioral dependent measures were collected during the immediate posttest. The behavioral measure was time on task during a two minute free play session; the alternate tasks were opportunity to read a popular magazine, simply sit quietly, or work on another Nina puzzle. The experimenter, who was blind to condition, "unobtrusively" recorded the amount of time the subjects spent on the puzzles. The questionnaire was then filled out, and included measures of intrinsic motivation of enjoyment, willingness to volunteer for future research, and desire for extra Nina puzzles. A delayed posttest, given one month later, included a third

Nina questionnaire designed to measure enjoyment of the task.

The measures of intrinsic motivation were found to be positively related. All four measures of intrinsic interest (volunteering, extra puzzles, time on task, and experimental enjoyment) were combined to form a single dependent variable.

A planned contrast, based on the assumption that the predicted effects were all equal in magnitude, was used to test the main effect of treatment. This contrast tested the hypotheses that: a) task-contingent rewards produce negative effects relative to conditions of no reward, b) positive feedback produces positive effects relative to conditions of no feedback, and c) these two effects should not interact. Additionally, Harackiewicz (1979) tested the hypotheses that performance contingent rewards decrease intrinsic motivation more than task contingent ones, and that performance-contingent rewards without norms supplied prior to task engagement decreased intrinsic motivation more than performance-contingent rewards with norms available at the outset. All hypotheses were supported statistically. In addition, it was reported that the positive feedback effect was stronger than the overjustification effect.

The Harackiewicz (1979) research also included some unique measures. A process variable, incidental recall,

was selected to measure the subjects ability to recall names and titles of the Nina puzzles. The amount of recall did not vary with the subject's initial enjoyment of the puzzles. As predicted, recall was lower in the two performance-contingent conditions. Harackiewicz (1979) predicted that subjects in the performance-contingent conditions would be more task oriented, and would recall fewer "irrelevant" details.

To check on the effectiveness of the positive feedback manipulation, perceived competency, (i.e., "How well do you think you did on these puzzles?" p. 1358) was measured. Subjects who received positive feedback perceived their performance to be more competent than those not receiving feedback, immediately after the experimental session and one month later.

In addition, the number of correct Ninas actually found on the three puzzles administered during the experimental session was recorded. All subjects found at least four Ninas, and no one found all fifteen. There were no significant effects for performance.

Harackiewicz (1979) concluded that:

> The control hypothesis received very strong support in this study. Performance-contingent rewards, particularly informational ones, were found to undermine subsequent intrinsic motivation more than task-contingent ones, which produced decrements in intrinsic motivation relative to control conditions of no reward... The enhancing effect of positive effect was found to be independent of the overjustification effect. (p. 1361)

Results reported by Harackiewicz (1979) differed from most previous research in two main respects. First, Harackiewicz (1979) found that a performance-contingent reward led to a greater loss of interest than a task-contingent reward. This is in direct contrast to previous studies (Boggiano & Ruble, 1979; Enzle & Ross, 1978; Karniol & Ross, 1977) which demonstrated that when standards of performance were incorporated into the experimental design, there were greater losses in subsequent intrinsic interest from task-contingent rewards than from performance-contingent rewards.

Second, Harackiewicz (1979) found that differences due to reward contingency were significant even in the presence of a socially comparative positive feedback information. In contrast, the Boggiano and Ruble (1979) study supported its prediction that, "If more direct information about competence or incompetence is additionally provided (in this case, performance level relative to others), subsequent level of interest should vary directly with the type of information regarding competence that is presented" (p. 1463). Specifically, Boggiano and Ruble (1979) found that for older children the reward contingency manipulation had no effect when comparative information was given; Harackiewicz (1979) found that the effects of the reward contingency

manipulation persisted, for older children, even when comparative social information was given.

Differences in results must take into account differences between the studies. There were no significant differences in the type of feedback manipulation, but there were differences as to when this feedback was presented. Harackiewicz (1979) utilized a normative, socially comparative standard, intertwined with an objective standard. However, the timing of presentation of this feedback differed from all previous studies in that it was presented only after the completion of the target task. In the Harackiewicz (1979) paradigm, most subjects have no access to any standards of performance, either objective or normative, until after completion of the target task. It is under these conditions that Harackiewicz (1979) found that reward contingency continued to operate despite competency feedback. It appears that when standards of performance are presented after the completion of the task, competency feelings may not be sufficiently increased to override reward contingency effects. Similarly, Harackiewicz (1979) found that a performance-contingent reward led to a greater loss of interest than a task-contingent reward, even when standards were presented. But these standards were not presented until after the completion of the task

engagement. In the one instance where positive feedback was provided prior to task engagement as well as afterwards, subsequent intrinsic interest did not decrease, relative to the control group (no feedback no reward).

Indirectly, (and possibly unrecognized by the author) Harackiewicz (1979) supported her contention that when "feedback is received earlier in the motivational process... (it) might have a greater positive effect than would feedback received after completion of the task" (p. 1354-5). Harackiewicz (1979) has demonstrated that positive feedback, provided after task engagement did increase feelings of competency. However, the positive effects of this feedback on competency feelings and subsequent level of intrinsic interest appear to be less powerful than when positive competency feedback is provided before task engagement.

The Harackiewicz (1979) study was unique in that the effects of tangible reward were separated experimentally from feedback, and the beneficial effects of positive feedback on subsequent interest level was clearly in evidence. Harackiewicz (1979) was the first researcher to directly measure the hypothesized intervening variable of competency feelings. The study revealed a direct relationship between positive feedback and an attitudinal measure of perceived competency performance. (Since

perceived competency was designed as a manipulation check, there was no attempt to measure the relationship between perceived competency and reward contingency.) The importance of positive feedback is emphasized by the Harackiewicz (1979) study, in that the positive feedback effect was reportedly stronger than the overjustification effect.

Rosenfield, Folger, and Adelman (1980) focused on the importance of rewards as a conveyer of information about competency. They attempted to explain discrepancies between the effects of task-contingent rewards and performance-contingent rewards (respectively termed "non-contingent" and "contingent" by Rosenfield et al. 1980). They maintained that, "It is not contingency per se, but rather whether the contingent rewards indicate substantial competence, that determines how the subjects will respond to these rewards" (p. 370).

To test this hypothesis, they manipulated the independent variables of level of pay (high or low) and reward contingency procedure (performance contingent/competency feedback; performance contingent/no feedback; task contingent/no feedback). The two additional conditions created were--no pay/competency feedback (high or low competence). Dependent measures were both behavioral and attitudinal. Orthogonal

contrasts among these eight conditions provided tests of the hypothesis.

Subjects were 118 female college students "who participated in the experiment to receive extra credit in their courses" (p. 370). The task was a modified version of the crossword game, "Ad Lib" (Milton Bradley). Students were presented index cards, each having 13 letters on it. The students were instructed to locate the corresponding letters on the dice, and then to proceed to arrange the crossword puzzle.

The experimental manipulation of feedback and competency were, by design, not orthogonal. Competency was manipulated by telling subjects in the performance-contingent group that they would receive a certain amount of pay for every letter combination they made. Theoretically, pay is conceived of as dependent on competency level. The task-contingent subjects were informed that they would be paid at the same hourly rate, regardless of the number of combinations they made. Theoretically, pay is independent of performance level.

Competency feedback was provided by telling all subjects that 70% of all students scored in the average range, 15% above average and 15% below average. In the competency feedback condition, subjects were told how well they did in comparison to others (experimentally assigned to high pay, above average, or low pay, below

average group). In the no feedback condition subjects were not provided with any comparative information concerning their performance.

High and low pay were conceived of as relatively high or low. In actuality all paid subjects received $1.75, but were told that others would earn $.25 (high pay condition) or that others would earn $5.00 (low pay condition). High and low pay were manipulated differently in each condition. And in the no pay condition, high and low pay reflected only low or high ability.

The dependent measures were assessed in an immediate posttest. A questionnaire was administered containing two measures of intrinsic interest: a) their impressions of the task in terms of interest, entertainment, enjoyability, and excitement; b) their willingness to come back at a later time to work on this task "for experimental credit only". They were asked to remain in the cubicle after filling out the questionnaire until the experimenter returned. The behavioral measure was amount of time spent playing on the Ad Lib game during this free time. There is no mention of alternative tasks available. Time on task, during a three minute period, was recorded by an observer.

Rosenfield, Folger and Adelman (1980) predicted that,

> When rewards for performing a task were an indication of competence on the task, higher reward would be associated with greater intrinsic motivation than lower rewards, but when rewards were not tied to competence--regardless of whether they were contingent--higher rewards would lead to lesser intrinsic motivation. (p. 374)

These predictions were confirmed. Moreover, students perception of their ability (as measured by their self rating of their skill as part of a manipulation check) supported the importance of perceived competency as a measurable intervening variable. Subjects in the performance competency/feedback condition, showed higher intrinsic interest on all three dependent measures, under high pay than under low pay, and competency feelings were higher. When rate of pay was determined randomly, there was no difference in feelings of competency, and high pay led to less intrinsic interest than low pay.

The authors concluded that:

> The presence or absence of competency feedback seems to be a crucial determinant of intrinsic motivation. When greater rewards indicate greater competence, increases in rewards are associated with increases in motivation. But when greater rewards do not mean greater competence, increases in rewards are associated with decreases in intrinsic motivation regardless of whether the rewards are contingent. (p. 374)

Rosenfield, Folger, and Adelman (1980) have demonstrated that competency feedback can be a significant factor in determining the effects of rewards on subsequent intrinsic interest. They manipulated

competency feelings directly by providing normative comparative feedback to subjects. Similarly to Boggiano and Ruble (1979), but unlike Harackiewicz (1979) they demonstrated that when comparative feedback is provided, interest level does not vary directly with the type of reward contingency. It is noted, that the Rosenfield et al. (1980) paradigm, provided feedback prior to task engagement. As noted by Harackiewicz (1979) this is earlier in the motivational process, and thus may account for its greater effects than found by Harackiewicz (1979) where feedback was provided after task engagement.

Rosenfield and his colleagues (1980) included a manipulation check of competency feelings. Thereby they demonstrated that feedback information can be manipulated, independent of contingency, to alter feelings of competency, as well as of intrinsic interest. Theoretically, they have succeeded in demonstrating that competency feelings are an important outcome of rewards that should be taken into account when predicting effects on intrinsic interest. From a practical point of view, it is questionable whether the effects of feedback and performance contingency can be separated. As noted by Harackiewicz (1979) in her discussion of appropriate control groups, "It can be argued that the feedback inherent in a performance-contingent reward is an integral part of the reward itself, in which case it

would not be unreasonable to compare performance-contingent rewards with task-contingent rewards or controls without positive feedback added" (p. 1362).

The Rosenfield et al. (1980) study, like the Harackiewicz (1979) study, served to focus attention on the importance of competency as an intervening variable between extrinsic rewards and subsequent intrinsic interest. Designed as theoretical treatises, their manipulations are contrived and do not resemble common behaviors in daily life. From a practical point of view they do not directly distinguish between those extrinsic rewards which enhance, and those that diminish subsequent intrinsic interest within an educational setting.

Karniol and Ross (1977), Enzle and Ross (1978), and Boggiano and Ruble (1979) demonstrated that a tangible reward could be contingently presented such that there was no loss in subsequent intrinsic interest, even when initial interest levels were high. Specifically, they demonstrated that a performance-contingent reward presented simultaneously with predefined attainable standards did not lead to a loss of intrinsic interest. They theorized that this was related to the ability of a reward to convey a message of competency. However, they did not directly measure the theorized changes in levels of perceived competency. Harackiewicz (1979) and

Rosenfield et al. (1980) further focused attention on the importance of competency an an intervening variable, and demonstrated that feelings of competency can be measured.

Chapter 3

Experimental Study

Rationale

The effects of tangible extrinsic rewards on intrinsic motivation remain controversial. There is a consensus of opinion that extrinsic rewards, "under certain circumstances" (Deci, 1975, p. 139), can produce a decrease in the subsequent level of intrinsic interest. These circumstances have not been fully delineated.

The focus on competency, as an intervening variable, appears to be the key to the understanding of the impact of rewards. Deci (1975) recognized that a person's intrinsic interest is enhanced when he feels competent at a task. Arkes (1978) proposed that a reward that causes feelings of competency will insulate behavior from the overjustification effect. The present study was designed to provide experimental support for these formulations.

Karniol and Ross (1977) maintained that when a reward conveys a clear message of competency, no loss of intrinsic interest follows. They demonstrated that when rewards were contingent upon successful completion of the task, and when they were based on obtaining specified and attainable standards of competency performance, there was no loss of subsequent intrinsic interest, even when

initial interest levels were high. Subsequent studies (Boggiano & Ruble, 1979; Enzle & Ross, 1978) replicated these findings. In all three instances only task-contingent rewards resulted in a loss of intrinsic interest; performance-contingent rewards caused no change in subsequent levels of intrinsic motivation. It is significant that these three studies, by design, included standards presented prior to task engagement, in a direct effort to arouse feelings of competency during task engagement. However, none of these studies measured the theorized changes in feelings of competency; none separated the effects of standards from reward contingency; none compared the effects of standards to conditions where no standards were provided; and, none allowed for direct comparison to earlier contradictory research results.

The present study was the first to include attitudinal measures of competency feelings as dependent variables, in order to evaluate statistically the effects of the experimental manipulation of reward contingency and standards on competency feelings. Harackiewicz (1979) and Rosenfield, Folger, and Adelman (1980) demonstrated that positive competency feedback can, independent of reward contingency, foster increased levels of intrinsic interest, and that competency feelings can be measured by questionnaires. However, in

the above studies, measures of competency feelings were designed as manipulation checks, rather than as dependent measures.

The present study incorporated three measures of perceived competency feelings. Personal competency was measured by childrens' responses to the question: "How do you think you did when you played this game?" Theoretically, provision of standards provides an external measure by which students can define their competency, and attainment of achievement at the prescribed standards level defines the individuals' performance as competent. Similarly, when rewards are presented as contingent on quality of performance, receipt of the reward serves to signal the individuals' performance as competent. Social competency was measured by childrens' responses to the following question, administered by the reward provider: "How do you think I would say you played this game?" Theoretically, receipt of a reward contingent on quality of performance serves to signal the individual that his/her competency is recognized by someone other than themselves, i.e., the reward provider. Task importance was measured by childrens' responses to the question: "How important do you think it is to do well on this game?" Task importance is conceived as a measure of the individual's perception of the importance of competent performance on

the targeted task. Theoretically, it has been recognized that, "Rewards which are allocated irrespective of the level of performance may indicate to the recipient that the quality of his performance is not important" (Karniol & Ross, 1977, p. 483). Additionally, rewards allocated depending on quality of performance should indicate to the recipient that the quality of his/her performance is important. And provision of standards should similarly indicate to the recipient that the quality of his/her performance is important.

The present study was the first to include a standards and no standards condition, and to orthogonally manipulate standards and reward contingency. This design allows for direct comparison of the seemingly contradictory findings reported by attribution researchers and operant psychologists. Attribution theorists have consistently reported a loss of intrinsic interest subsequent to the introduction of tangible rewards. Attribution theorists have typically selected paradigms in which rewards were presented for engaging in a task, without providing objectively defined standards of competency. It is maintained that under such conditions reward does not convey a message of competency. Behaviorists have consistently reported no loss of intrinsic interest subsequent to the introduction of tangible rewards. Behaviorists, have customarily

utilized rewards as performance-contingent, with clear objective standards which are attainable by the subject and are presented prior to task engagement. It is maintained that under such circumstances competency feelings are engendered. It is contended that increased competency feelings result in increased levels of intrinsic interest and serve to insulate behavior from the overjustification effect.

On the basis of past experimental research, it is hypothesized that after being given achievable standards, students receiving performance-contingent rewards will increase their feelings of competency and intrinsic interest compared to students given task-contingent rewards when no standards are provided. Perceived feelings of personal competency, social competency, and task importance are conceived as intervening variables which summate to affect subsequent levels of intrinsic interest. In order to test these hypotheses, a research design that involves variations in standards and reward contingency, and that measures differences in perceived competency feelings as well as intrinsic motivation, was utilized.

Hypotheses

With regard to the student's intrinsic interest, it was hypothesized that:

Hypothesis 1: Intrinsic interest is greater under conditions of performance-contingent reward and standards than under task-contingent reward and no standards.

Hypothesis 2: Intrinsic interest is less under conditions of task-contingent reward and no standards, than under no reward and no standards, a control condition.

Hypothesis 3: Intrinsic interest is greater under conditions of performance-contingent reward and standards, than under no reward and no standards, a control condition.

With regard to students' perceived personal competence, it was hypothesized that:

Hypothesis 4: Perceived personal competence is greater under standards conditions than under no standards conditions.

Hypothesis 5: Perceived personal competence is greatest under conditions of a performance-contingent reward. Specifically, a) perceived personal competence is greater under conditions of a performance-contingent reward than under no reward; and b) perceived personal competence is greater under conditions of a performance-contingent reward than under a task-contingent reward.

Hypothesis 6: Perceived personal competence is greater under conditions of performance-contingent reward

and standards, than under task-contingent reward and no standards.

Hypothesis 7: Perceived personal competence is greater under conditions of performance-contingent reward and standards, than under no reward and no standards, a control condition.

With regard to the students' perceived social competence it was hypothesized that:

Hypothesis 8: Perceived social competence is greatest under conditions of a performance-contingent reward. Specifically, a) perceived social competence is greater under no reward; and b) perceived social competence is greater under conditions of a performance-contingent reward than under a task-contingent reward.

With regard to the students' perceived task importance, it was hypothesized that:

Hypothesis 9: Perceived task importance is greater under standards conditions, than under no standards conditions.

Hypothesis 10: Perceived task importance is greatest under conditions of performance-contingent rewards, and least under conditions of task-contingent rewards. Specifically, a) perceived task importance is greater under conditions of a performance-contingent reward than under no reward; and b) perceived task

importance is greater under conditions of a performance-contingent reward than under a task-contingent reward; and c) perceived task importance is greater under conditions of no reward than under a task-contingent reward.

Hypothesis 11: Perceived task importance is greater under conditions of performance-contingent reward and standards, than under task-contingent reward and no standards.

Hypothesis 12: Perceived task importance is greater under conditions of performance-contingent reward and standards, than under no reward and no standards, a control condition.

Pilot Testing

Pilot testing was conducted to develop and ascertain the appropriateness of experimental materials, specifically, the target task, the reward, and alternate experimental materials.

Target task. Pilot testing was conducted to develop an appropriate target task. The target task was designed as two varied but similar drawings of the movie character "E.T.". One version was printed on yellow paper and one on white paper (see Appendixes B & C). The students were instructed to mark as many of the differing details as they could find. Pilot testing was conducted a) to determine appropriate time allotment for the

experimental target task, and b) to determine the standard of performance that would likely be attainable to all students, and c) to ascertain that the level of initial interest in the target task was sufficiently high.

Eighty-six students in second, third and fourth grade classes were presented with the target task by a cluster teacher who routinely taught each of the classes. Students were presented with the target task without any reference to rewards or standards of good performance. a) In order to determine the appropriate time allotment for the experimental target task, students were given unlimited time to complete the task during the pilot testing. b) In order to determine the standard of performance which would likely be attainable to all students, the completed target tasks were scored to provide a measure to task difficulty. c) In order to ascertain that the level of initial interest in the target task was sufficiently high, students were provided with a free play period during which they had the opportunity of working on additional puzzle activities of the same type as the target task.

Time on the target task during pilot testing was highly variable and ranged from two to ten minutes. Most of the students handed in their drawings within the first five minutes. The teacher reported that almost all of

the students stopped working after the first five minutes, although many held on to their papers beyond that time period. Consequently, for the experimental study, time on target task during experimental manipulation was set as five minutes.

Grading of the target task indicated that out of a total of eleven possible differences between the two pictures the average number of errors noted was 5.5, with individual scores ranging from three to nine errors detected. There was no direct correspondence between grade level and scores obtained (see Table 3). Consequently, for the experimental study, the standard of good performance for all students was set as three or more errors detected.

The level of initial interest in the target task, during pilot testing, was found to be very high. During the free play period immediately following the target task assignment, 80 out of 86 students (93%) voluntarily selected to keep busy with "some more drawings like the ones you have just finished". Children in all three grades demonstrated strong interest in the target task, with 93% of the second graders, 100% of the third graders, and 86% of the fourth graders requesting additional drawings.

Reward. Pilot testing was independently conducted to ascertain the appropriateness and relative

desirability of the reward. A "Good Player Award" certificate with a personal picture of the student, served as the experimental reward.

Second, third, and fourth grade classes were addressed individually by a cluster teacher who regularly taught all three classes. Students were asked to help a teacher in another school select the most desirable award to give for excellence on a science project. The three awards displayed were Hershey's Candy Kisses, two new pencils, and the "Good Player Award". Students were instructed to vote for the award that they thought other children "will like best". In a further check for preference, each possible combination of two awards was directly compared, and students were again instructed to vote for the award that they thought other children "will like best".

Given a choice of three possible awards, 74% of the children chose the certificate, 13% chose the pencils, and 13% chose the candy. Preference for the pictured certificate exceeded 50% in each of the three grades (see Table 1). When each possible combination of two awards was directly compared, preference for the picture certificate was again strong. Preference for the "Good Player Award" exceeded 70% in each of the grades, with 83% of the students preferring the certificate to the pencils, and 86% preferring the certificate to the candy

(see Table 2). Consequently, the "Good Player Award" was considered to be appropriate and highly desirable for primary school students, and was utilized within the experimental study as the reward.

Alternate tasks. Informal testing was conducted to determine the appropriateness of the alternate activities to be made available to students during the experimental free play period. A total of thirty children from second, third and fourth grades worked individually with the experimenter. The three planned alternate activities were find-the-hidden-word, reading material, and a drawing activity. Children were presented with variations of the tasks, and were instructed that they could play with any or all of the available activities for as long as they wished. It was found that all of the activities were of interest to the students of all three grades, and materials suitable for primary grade students were selected. However, one task, find-the-hidden-word, appeared to have an implicit level of standards, such that many children appeared unhappy with their performance, and some actually requested standards or verbal reassurances. This task was therefore omitted in the experimental design. Based on this pilot testing the amount of time that all students maintained continuous interest in the activities was set as under fifteen minutes. This pilot testing also served

to determine the number of materials required for each activity, so that students could conceivably spend the entire free play time on any one of the tasks.

Methodology

Subjects. Subjects were 96 second, third, and fourth grade boys and girls, who lived and attended school in Long Island, New York. The children were randomly assigned to experimental conditions. Without exception, all followed the outlined procedures from the outset to the completion.

Materials: The target task consisted of two varied but similar pages from a coloring book, one printed on a white background and one on a yellow background. The picture on the yellow page had missing or different details from those on the white page. The task presented to all the children, was to circle as many things as they could which were on the yellow page but were either missing or different on the white page. A black marking pen was provided.

The reward was designed to be both symbolic and tangible. Both behavioral and attributional researchers have used such rewards previously in their investigations. Specifically, an award certificate with a personal picture of the subject served as the reward. The certificate was printed on golden heavy weight construction paper that was 8 by 11 inches. It was

decorated with a black design, was inscribed as a "Good Player Award", and had place for the children's name and photograph. The photos were standard sized colored snapshots taken with a 110 pocket camera. The pictures were taken immediately after experimental manipulation or after completion of all data collection, depending on the experimental condition. In the presence of the student, their name was inscribed on the award certificate. It was explained that the award certificate would be delivered after the roll of film was developed. Reading material and fill-in drawing tasks served as alternate activities during the free play period. Neither of these activities included any explicit standards of achievement.

Instructions during experimental manipulation were presented via a tape recorder and earphones. Six different cassette tapes were available, one for each condition. The tapes were coded for reward and no reward conditions. Subsequently they were numbered one through six. The experimenter was informed only as to the necessity of providing or not offering a reward and thus remained blind concerning other experimental conditions during the entire experimental procedure.

Design. The research design was a 3 X 2 factorial involving type of reward and task standards (see Figure 1). The reward manipulation consisted of a no-reward

group, a task-contingent reward group, and a performance-contingent reward group. Task contingency was defined as an expected reward, contingent only upon task engagement. Performance-contingent reward was defined as an expected reward that was contingent upon successful performance on the target task.

The task standards manipulation involved a no standards and a standards condition. No external measure of competency evaluation was given to the no standards group. The standards group was provided with clearly defined external standards, non-normative, and predetermined to be attainable to the sample population. Standards were provided prior to task engagement.

The performance-contingent reward with standards condition was assumed to be the most similar to the paradigm utilized in behavior modification programs. The task-contingent reward with standards condition was perceived as most closely approximating the basic paradigm utilized by attribution theorists. The no reward, no standards condition was perceived as the control group.

The dependent measures consisted of a behavioral scale and three attitudinal scales. Intrinsic interest was behaviorally measured by time on target task during free play. Three measures of task attitude were assessed by self-report questionnaires, with data collected prior

to free play. The attitudes measures included: perceived personal competence; perceived social competence; and perceived task importance (see Appendix A). These attitude measures were conceived as intervening variables that correlated directly with intrinsic interest, measured by time on target task during free play.

Procedures. The children were introduced to the experimenter in small groups. They were told that they would have the opportunity to come to the "Surprise Room" individually. After arriving in a separate experimental room, the children were seated by a tape recorder, earphones, and a target task. A sample award certificate hung in front of the room, but no reference was made of it. The children were helped to adjust their earphones and were informed that they should follow the instructions on the tape. (Appropriate cassettes were previously inserted into the recorders according to experimental condition.)

All the children heard the following on the cassette tape:

> "Hello. I hope you can hear my voice clearly. If the sound is too soft or too loud raise your hand so that it can be made better.... In front of you there are two pictures. These pictures are almost the same, but there are some things on the yellow page which are either missing or different from the white page. When I tell you to start, you are to use the black magic marker to circle as many things as you can find on the yellow page that are either missing or different from the white page. Wait until

I tell you to start and then make your circles on the yellow page."

Instructions were then varied according to randomly assigned condition. (See Appendix D for differential instructions grouped according to experimental condition.)

Children assigned to the standards condition were told, "If you can find and circle three or more things, you will be doing a good job." Children assigned to the performance-contingent reward condition were told, "For a good job I have a special prize for you. It is a golden award certificate with your own picture on it. It is like the award certificate hanging in front of this room." Children assigned to the task-contingent reward condition were told, "For playing this game I have a special prize for you. It is a golden award certificate with your own picture on it. It is like the award certificate hanging in front of this room."

All subjects were told, "You will have enough time. Please stop when the buzzer rings." Children assigned to the performance-contingent reward condition were told, "Then if you have done a good job, I will make out your prize certificate." Children assigned to the task-contingent reward condition were told, "Then I will make out your prize certificate." All subjects were told, "The instructions will now be repeated"

(appropriately for each cell). "You are ready to start. Take off your earphones and go ahead."

After five minutes elapsed, the buzzer sounded. The experimenter removed the tape from the recorder and the children were thanked. No reference was made to achievement on the task. Subjects who were in one of the reward conditions were immediately shown their award certificates, their name was written in, and their pictures were taken. The attitudinal measures of perceived personal competence, social competence, and task importance, were then administered to the children.

The children were introduced then to each of the three tasks in the free play area (variations of target task and two alternate activities). The youngsters were informed that they could play with any of the three activities in the room. All were told that, "There will be no grading, no awards, no prizes. You can play with anything you like." The examiner then seated herself behind the free play area so that she could record time-on-target-task unobtrusively. On-task behavior on the target task was defined as holding or looking at the E.T. task, as well as actually marking the picture. The free play period lasted twelve minutes.

After the free play period, children who were assigned to the no-reward experimental conditions were shown a reward certificate, their name was filled in,

their picture taken, and they too were informed that the award would be delivered after the film was developed. All subjects were reminded that, "This is a surprise room", and were encouraged to help keep it a surprise for the other children. Questions by the children were deferred and all were informed that their questions would be answered after all other children had their chance to play.

Debriefing of subjects was conducted after the collection of the data was completed and after the award certificates were distributed. Debriefing included informal group discussion with the children as well as answering questions.

Chapter 4

Results

The design involved two factors with three levels of reward manipulation (performance-contingent reward, task-contingent reward and no reward) and two levels of standards (standards and no standards). The dependent measures included a measure of intrinsic interest and three measures of task attitude.

The first three hypotheses of the study were focused on the behavioral variable of level of intrinsic interest. This was measured using time on target task during free play, and the results for each of the six experimentally manipulated groups are presented in Table 4.

Three a priori contrasts using one-tailed t tests, were conducted to directly test the first three hypotheses of the study. The first hypothesis, predicted greater intrinsic interest under conditions of performance-contingent reward and standards, than under task-contingent reward and no standards. It was supported statistically, t (30) = 2.15, p <.05. During free play, the group receiving performance-contingent reward and standards spent 270 seconds on the target task, and the group receiving task-contingent reward and no standards spent only 154

seconds on the target task. This was the greatest difference between any of the six groups.

The second hypothesis predicted less intrinsic interest under conditions of task-contingent reward and no standards than under conditions of no reward and no standards. Group means show a trend in this direction (see Table 4), but the difference was not statistically reliable.

The third hypothesis predicted greater intrinsic interest under conditions of performance-contingent reward and standards than under no reward and no standards. Group means show a trend in this direction (see Table 4), but the difference was not statistically significant.

The three attitudinal dependent variables were assessed by questionnaires: perceived personal competence, perceived social competence, and perceived task importance. Scores on the scales ranged between one and five, with a score of five representing feelings of greatest competency or task importance.

The mean scaled score of perceived personal competence for each of the six experimentally manipulated groups is presented in Table 7. The overall mean score in perceived personal competence was 3.98. This approximates a rating of "very good" (see Appendix A).

Five a priori contrasts utilizing one-tailed t tests, were conducted to statistically test for the significance of the predicted effects on perceived personal competence as specified in hypotheses four through seven.

It was predicted that perceived personal competence is greater when standards are given. This was statistically supported, t (94) = 1.93, $p < .05$. Children given standards showed higher feelings of perceived personal competence than children who were not presented with standards (see Table 7).

The prediction that perceived personal competence is greatest under conditions of a performance-contingent reward was only partially supported. The prediction that perceived personal competence is greater under conditions of a performance-contingent reward than under a task-contingent reward was statistically supported, t (62) = 1.92, $p < .05$. Under performance-contingent conditions, children showed higher feelings of perceived personal competence than children who received a reward under task-contingent conditions (see Table 7). However, the prediction that perceived personal competence is greater under conditions of a performance-contingent reward than under no reward was not supported statistically. Inspection of group means indicates that although differences were found in the predicted

direction, the differences were not statistically reliable (see Table 7).

Two predictions were made concerning the effects of specific experimental manipulation on perceived personal competence. The prediction that perceived personal competence is greater under conditions of performance-contingent reward and standards compared to task-contingent reward and no standards was supported statistically, t (30) = 3.13, $p < .05$. Under conditions of performance-contingent rewards and standards, the mean score on the perceived personal competence scale was 4.25; under conditions of task-contingent reward and no standards, the mean score in perceived personal competence was 3.38. This reflects the greatest contrast in scores for any of the six experimental groups in perceived personal competence (see Table 7). The prediction that perceived personal competence is greater under conditions of performance-contingent reward and standards than under no reward and no standards was not statistically supported.

Perceived social competence was conceived as an intervening variable that would be enhanced by the presentation of a performance-contingent reward. A priori comparisons, using one-tailed t tests, failed to provide statistical support for these predictions. Perceived social competence was not statistically greater

under conditions of a performance-contingent reward than under no reward. Nor was perceived social competence greater under conditions of a performance-contingent reward than under a task-contingent reward. Analysis of the mean scores indicated that these differences due to the experimental manipulation, although in the predicted direction, were generally small. The most notable difference between the groups was in the low rating of the group that received task-contingent reward and no standards (mean = 3.63), and all the other groups. All other group ratings ranged from 3.94 to 4.13. The overall mean score in perceived social competence was 3.97 which approximates a rating of "very good" (see Table 9, & Appendix A).

Six a priori contrasts, using one-tailed $\underline{t}$ tests, were conducted to test for the predicted effects on perceived task importance as specified in hypotheses nine through twelve. There was no statistical support for any of the hypotheses regarding perceived task importance.

Mean scores obtained on the scale of perceived task importance for each of the six experimentally manipulated groups is presented in Table 11. The overall mean rating of task importance was 2.98, which is almost exactly in the center of the allowable choices. This approximates a rating of "important" (see Appendix A).

Five separate analyses of variance were conducted. The independent variables were standards (standards and no standards) and contingency of reward (performance-contingent reward, task-contingent reward and no reward). The dependent variables were a behavioral measure of intrinsic interest, three measures of task attitude, and a measure of performance on the target task. Because no gender differences were predicted and because preliminary analysis of data by sex of subject revealed no significant gender differences, t (94) = 0.56, $p > .50$, the data were collapsed across this dimension.

There were no significant effects of standards or reward contingency on subsequent intrinsic interest on the target task, F (5,90) = 1.43, $p > .20$, (see Table 5). (A time transformation was conducted to see if it could improve the amenability of the data to parametric analysis. However, the results proved to be statistically comparable. Therefore, it was decided to report the raw data, given its more straightforward interpretation.) There were marginally significant effects of experimental treatments on perceived personal competence, F (5,90) = 2.19 $p = .06$. Students in the standards conditions showed higher feeling of perceived personal competence than students in the no standards condition. (see Table 6). There were no significant

effects of experimental treatments on perceived social competence, F (5,90) = 0.76, p > .50, (see Table 8). There were no significant effects of experimental treatments on perceived task importance, F (5,90) = 0.93, p >.40, (see Table 10).

No predictions were made regarding the effect of standards or type of contingency performance on the target task. However, since subjects' scores varied highly, the possibility that experimental manipulation may have affected performance on the target task was explored. Analysis of variance, yielded no statistically significant effects, F (5,90) = 1.26, p > .20, (see Table 12). The overall mean number of errors detected by the subjects was 5.39 (see Table 13).

Pearson product-moment correlations for pairs of variables were conducted for the behavioral measure of intrinsic interest and the three self report measures of task attitude (See Table 14). The correlations of the behavioral measure of intrinsic interest (time on task during free play) and the attitudinal measures of perceived personal competence, r = .24, were statistically reliable. The correlation's of the behavioral measure of intrinsic interest and the attitudinal measure of perceived task importance, r = .26, were statistically reliable. The three attitudinal measures intercorrelated significantly, although the

proportion of variance explained was small. Perceived personal competence correlated with perceived social competence, r = .38, and with perceived task importance, r = .21, and perceived social competence correlated with perceived task importance, r = .25, (see Table 14).

Chapter 5

Discussion

Theoretical Implications

Performance contingency and standards. This study was designed to determine some of the conditions that differentiate between positive and negative effects of reward presentation. Prior research had demonstrated that the effects of extrinsic rewards on subsequent intrinsic interest were dependent upon the conditions under which the rewards were presented. To date the effects of standards had not been separated from the type of contingency. In the present study, performance-contingent rewards that were presented with attainable standards of performance created a higher level of subsequent intrinsic interest than when rewards were presented as task-contingent without any standards of performance.

The experimental results indicated that the differences in conditions of presenting extrinsic rewards in behavior modification programs and in attributional research paradigms produced different levels of subsequent intrinsic interest. Presentation of a reward contingent upon successful performance, presenting clearly defined standards prior to task engagement, and setting standards that are attainable, is the paradigm which customarily reflects conditions utilized in

behavior modification programs. In contrast, presenting rewards contingent upon task engagement, with no specified standards of performance, is the original paradigm utilized in attribution oriented research. It was found that interest levels, as measured by time on target task during free play, remained significantly higher when conditions replicated the behavior modification paradigm than when they followed the attributional research paradigm.

The present study was designed to experimentally separate the effects of providing standards and the effects of reward contingency on subsequent levels of intrinsic interest. Analysis of variance yielded no statistically significant main effects for either reward contingency or standards, on levels of intrinsic interest (see Table 5). Since differences between the experimental groups on time on task behavior appear to indicate large directional trends (see Table 4), results are analyzed in combination with previously reported research.

In this study a task-contingent reward consistently depressed students' levels of intrinsic interest, relative to the non-rewarded group, although this difference did not reach statistical significance, $p < .12$ (see Table 5). Previous research, however, lends support to the contention that these differences were related to reward contingency. Results were consistent

with reported findings that a task-contingent reward reduced subsequent levels of intrinsic interest both under conditions where standards were provided (Boggiano & Ruble, 1979; Enzle & Ross, 1978; Karniol & Ross, 1977), as well as when no explicit standards were provided (Lepper, et,al., 1973; Harackiewicz, 1979).

The present study is interpreted as supporting the contention by Harackiewicz (1979) and Rosenfield, Folger, and Adelman (1980) that reward contingency cannot independently predict subsequent intrinsic interest levels. The results of this study indicate that when explicit and attainable standards are not provided, the effects of a performance-contingent reward relative to no-reward, remains undeterminable. Greene and Lepper (1974) reported that a receipt of a reward contingent upon "best" performance resulted in decreases in subsequent levels of intrinsic interest. In the present study, under similar conditions (performance-contingent reward and no standards), no decrease resulted. Contradictory results are probably related to the ambiguousness of the instructions. When "good" or "best" is left undefined, the nature of the task or the age of the children may conceivably become important variables, affecting competency feeling and intrinsic interest levels. It is therefore contended that the receipt of a performance-contingent reward cannot be depended upon to

maintain levels of intrinsic interest, and that the simultaneous provision of a performance-contingent reward and standards is required in order to consistently predict maintenance of intrinsic interest.

This is the first study to orthogonally manipulate reward contingency and standards. No main effects of standards on intrinsic interest were predicted and no statistically significant effects were found (see Table 5). Inspection of differences between the experimental groups (see Table 4) indicates that providing standards may be an effective means of increasing levels of intrinsic interest, especially when no reward is involved. Although this is speculative at present, the results are sufficiently significant (see Table 5) to warrant further research.

It is emphasized that standards, as utilized in the current study, were predefined as attainable and as presented prior to task engagement. Standards which are either not attainable or which are not presented prior to task engagement are not expected to be effective. Karniol and Ross (1977) reported that providing subjects with standards they couldn't attain resulted in less intrinsic motivation. Harackiewicz (1979) reported lower intrinsic motivation when standards were presented prior to task engagement compared to after task engagement.

One unfortunate shortcoming of the present attempt to select standards that were attainable by all subjects was that such standards did not optimize task performance. Analysis of variance indicated no effects of either standards or reward contingency on subjects score on the targeted activity (see Table 12). Comparisons of the scores obtained by children in the standards group, with those in the no standards group and those in the pilot group (where no standards were provided), indicated that the average score was always between five and six (see Tables 3 & 13).

In the present study, the presentation of low standards of performance prior to task engagement, did not affect subsequent levels of performance on the target task. Nevertheless, caution should be utilized when presenting artificially low standards. If standards had served to reduce performance levels, subsequent intrinsic interest might have been lowered by the use of standards. It is noted that the standards provided in this study indicated that a response of three or more would be considered as "good". The importance of adding the terminology "or more" may have been a key factor in eliminating a loss of motivation. In the present study the utilization of artificially low standards may have served to limit the extent to which competency feelings were increased. It is possible that utilization of

higher, more realistic, and more challenging standards may prove more effective in increasing subsequent performance levels, competency feelings, and subsequent motivational levels. There remains a need for further research to study the manner in which standards should be selected.

Competency feelings. This is the first study to directly examine the effects of manipulation of standards and reward contingency on feelings of perceived competency. The results support the conclusion that perceived feelings of personal competency is a key intervening variable which is directly related to levels of intrinsic interest in a prescribed task. Pearson product moment correlations indicated that rating of personal competency correlated significantly with time on task (see Table 14).

In this study, perceived personal competence was measured immediately after the experimental manipulation. Children responded to the question, "How well do you think you did when playing this game?", with one of five given responses. The highest score in perceived personal competence was within the group receiving both a performance contingent reward and standards. It was concluded that the increase in feelings of perceived personal competence resulted from the additive effects of presenting the reward as performance contingent, and the

provision of standards. As predicted, providing attainable standards to the subjects prior to task engagement resulted in a consistent and statistically significant increase in feeling of personal competence. It was concluded that providing standards prior to task engagement allowed children to experience success during their original encounter with the target task, and this in turn increased their feelings of personal competence. Similarly, as predicted, presenting rewards contingent upon "good" performance (defined or undefined), compared to presenting them contingent on task engagement resulted in a consistent and statistically significant increase in feelings of personal competence. It is suggested that when subjects received rewards because they did a "good job" the rewards conveyed additional information that they had performed successfully.

In this study, the largest contrast in level of perceived personal competence occurred between children who received both standards and a performance contingent reward and those who received task contingent rewards and no standards. (see Table 7). Similarly, the largest contrast in level of intrinsic interest occurred between children who received both standards and a performance contingent reward and those who received task contingent rewards and no standards (see Table 4). Both these comparisons were statistically significant. This

confirms the conceptualization of perceived personal competence as an intervening variable which can explain differences in subsequent interest levels when extrinsic rewards are utilized in behavior modification programs and when they are utilized in attributional studies. Furthermore, it supports the contention by Arkes (1978) that feelings of competence will insulate behavior from the overjustification effect. Theoretically, the different effects of rewards on subsequents levels of intrinsic interest, appears to be related to the extent to which the total mode of reward presentation resulted in an increased perception by the individual that they were personally competent at the targeted task.

It was predicted that experimental manipulation of reward contingency and standards would result in changes in feelings of social competence and in feelings about task importance. In this study, perceived social competence and task importance were measured after the experimental manipulation and immediately following the measure of personal competency. Children sequentially responded to the question, "How do you think I would say you played this game?" and "How important do you think it is to do well on this game?", with one of five given responses. Although feelings of social competency and task importance were related to each other and to

perceptions of personal competency (see Table 14) there was no direct effect of experimental manipulation.

A priori comparisons and analysis of variance failed to produce statistical support for any hypothesized changes or for any statistically significant main effects of experimental manipulation on feelings of social competence (see Table 8) or perceived task importance (see Table 10). Nevertheless, there are two contrasts which are noteworthy. Children who received task-contingent rewards without standards scored the lowest relative to all groups of children, in both perceived social competence and task importance. The validity of the conceptualization of social competency and task importance remains theoretically interesting but experimentally unsubstantiated.

The primary intervening variable appears to be the perceived feelings of personal competence. Further research is needed to develop more reliable and valid measures of personal competence.

Non-significant findings. This study did not replicate earlier findings by attributional theorists that task contingent rewards reduced levels of intrinsic interest relative to control groups. Lepper and Greene (1973) reported that under conditions where no explicit standards were provided, a task contingent reward produced less task interest than displayed by children in

a control group. However, in the current study, a priori comparison of level of intrinsic interest for children who received a task contingent reward and no standards and youngsters who received no reward and no standards (control group) was not significant. Analysis of variance procedures showed no main effect for reward contingency (see Table 8). It is noted that although previous findings were not statistically replicated, differences in the subsequent level of intrinsic interest were in the predicted direction.

One possible explanation for this lack of replication could have been due to a lack of initially high interest in the target task. However, pilot testing indicated that the task was of high initial interest. Within a free play situation, 93% of the students spent time on activities similar to the target task. Furthermore, the large amount of time that students spent on the targeted task during free play also reflects a high level of initial interest in the target task. The control group spent an average of over three minutes on the target task, out of a possible total of twelve minutes (see Table 4). Considering that each youngster had spent a full five minutes on the target task immediately prior to the free play period, this is a high proportion of time devoted to the target task despite the availability of two to other interesting tasks. The

childrens relatively high level of initial intrinsic interest in the target task was displayed during both pilot and experimental testing.

Lack of replication of previous attributional research may be attributable to the generally high feelings of competence in the target task. On the self report questionnaire of perceived personal competence the average response of all subjects and the average response of the control group approximated a rating of "very good" (see Table 7). Given five choices to the question, "How do you think you did when you played this game?" responses indicated that the subjects perceived themselves as highly competent on the target task, even when no standards were presented. Most children rated their performance as good, very good, and excellent; only three rated it as fair, and none rated it as poor. Again, it appears that when youngsters perceive themselves as competent on a task, any loss of intrinsic interest that may otherwise be induced by reward, is counteracted by those feelings.

This explanation and interpretation of the data is offered merely as tentative since comparative data concerning the competency feelings of subjects toward the targeted task in prior studies is not available. Earlier researchers did not measure perceived feelings of competence. Further research is recommended wherein

perceived competency levels in the target task are predetermined to be low. Replication of current results, while simultaneously replicating previous attribution research (where task-contingent rewards reduced levels of intrinsic interest relative to control groups), would provide further confirmation of the competency hypothesis.

The present study was unable to support the prediction that level of intrinsic interest would be greater under conditions of a performance-contingent reward and standards than under conditions of no reward and no standards (control group). An a priori contrast indicated that although the trend was in the predicted direction, the differences were not statistically significant.

Theoretically, an increase in level of intrinsic interest in the group receiving both a performance-contingent reward and standards was hypothesized, based on the additive effects of predicted gains in perceived personal competence, social competence, and task importance. However, when comparing students who received performance contingent rewards and standards to the control group, only manipulation of standards raised competency feelings, and this was limited to feelings of personal competence. Evidently,this effect was insufficient to independently

raise levels of intrinsic interest. Consequently, there was no significant difference in levels of subsequent intrinsic interest between children who received reward in the mode of behavior modification programs and youngsters who received neither reward nor standards.

It is important to emphasize that the conditions utilized in this study, even in the performance contingent reward and standards group, were not exact replications of a behavior modification program. Since this was a one trial paradigm and since the free play session may be viewed as an extinction period, positive results would not necessarily be predicted, by even proponents of behavior modification theory.

Notable Design Features.

There were three aspects of the experimental design which warrant consideration when designing further research. In the present study, children received the major part of their instructions via a tape recorder and earphones. The original intent was two fold: a) to eliminate variation in instructions (wording, voice quality, body language) both within and between experimental groups; and b) to allow one experimenter to conduct the entire procedure, including the experimental manipulation, while remaining blind to the experimental condition being used. These goals were met. The use of

the tape recorder with earphones was additionally helpful in eliminating extraneous noises.

The reward presented to the students was original in design and method of distribution. The award certificate was individualized with both the name and picture of the student. The use of a personalized photo had several advantages. This award certificate was valued and desired by the youngsters in grades two, three, and four. The use of the photo also made it convenient to distribute the awards at a future date. The children readily accepted the explanation that their receipt of the award would be delayed until the film could be developed. This insured that students never left the experimental room with any tangible evidence of their participation in the experimental study. This may have served to reduce unwanted discussion among students about the experimental session while simultaniously protecting the value of the award certificate. Saliency of the award was insured as youngsters were immediately able to see their individual award certificate with their names inscribed.

The use of particular terminology, describing the experimental room as a "surprise room", was unexpectedly found to be advantageous with elementary school children. The title was selected in an effort to minimize questions prior to experimental manipulation. However, its

greatest asset appeared to be in its effect on silencing inter-subject discussion. The children seemed to take the room title seriously. During the debriefing it became evident that discussion among the students about the activities in the "surprise room" had been minimal. Only five students reported some foreknowledge about the target task (i.e., "E.T." characters). None reported knowledge about the reward or picture taking procedures prior to their actually entering the surprise room. Most interestingly, the students repeatedly voiced their opinion that is would not have been fair or nice to spoil the surprise for the other children. The simple use of the terminology "surprise room" was apparently helpful in controlling contamination of results by eliminating a possible pre-experimental expectation of a reward.

Educational Implications.

The results of the study have direct implications for classroom application. They will serve to reassure educators about their use of behavior modification procedures. Fears that extrinsic rewards will decrease subsequent levels of intrinsic interest, engendered by the early attribution research, can be allayed. It has been demonstrated that the adverse effects of an extrinsic reward on subsequent intrinsic interest, are limited to those procedures used by attribution researchers. Teachers can be reassured that a properly

administered behavior modification program need not result in a loss of intrinsic interest, even among those students who showed an initially high level of interest in the target task.

Specifically, the research study provides theoretical and practical guidelines for teachers when they use extrinsic rewards, whether or not a behavior modification program is implemented. Practically, it has been demonstrated that extrinsic rewards should be presented as dependent upon successful performance rather than for simply engaging in a targeted task. And it emphasized the simultaneous need for clearly specified standards of performance which are designed as attainable and are presented prior to task engagement. Theoretically, the results indicate that the overriding factor is the need to present rewards in a manner which fosters feelings of personal competence.

Table 1

Pilot Testing:

Reward Preference Scale - Three Way Comparison

	2nd grade	3rd grade	4th grade	Total
Pencils	17.9%	8.9%	12.9%	12.9%
Certificate	53.5%	82.3%	83.9%	74.2%
Candy	28.6%	8.8%	3.2%	12.9%

Table 2

Pilot Testing:

Reward Preference Scale - Two Way Comparison

	2nd grade	3rd grade	4th grade	Total
Pencils	28.6%	11.8%	3.0%	17.0%
Certificate	71.4%	88.2%	87.0%	83.0%

Table 3

Pilot Testing: Scores On Target Task

	2nd grade (27 S's)	3rd grade (30 S's)	4th grade (29 S's)	Total (86 S's)
Mean	5.6	5.8	5.6	5.7
Mode	5.5	6	5	5.5
Median	6	6	5	6
Range	3-8	4-9	4-8	3-9

Table 4

Mean Time On Target Task During Free Play
(Reported in seconds)

		Standards	No Standards	Total
Performance Contingent Reward				
	M	269.69	205.75	237.72
	SD	190	177	
Task Contingent Reward				
	M	157.50	153.63	155.56
	SD	162	101	
No Reward				
	M	263.06	187.94	225.50
	SD	198	164	
Total				
	M	230.08	182.44	206.26

Table 5

Analysis Of Variance Of Time On Target Task On Rewards and Standards
(N=96)

Source	df	MS	F	p
Main Effects	3	60082	2.106	0.105
Rewards	2	62881	2.205	0.116
Standards	1	54483	1.910	0.170
Interaction	2	11746	0.412	0.664
Explained	5	40747	1.429	0.222
Residual	90	28523		

Table 6

Analysis Of Variance Of Perceived Competency Personal On Rewards and Standards (N=96)

Source	df	MS	F	p
Main Effects	3	1.812	2.615	0.056
Rewards	2	1.385	1.999	0.141
Standards	1	2.667	3.848	0.053
Interaction	2	1.073	1.548	0.218
Explained	5	1.517	2.188	0.062
Residual	90	0.693		

Table 7

Mean Score On Perceived Competency Personal Scale

	Standards	No Standards	Total
Performance Contingent Reward	4.25	4.06	4.16
Task Contingent	4.13	3.38	3.75
No Reward	4.06	4.00	4.03
Total	4.15	3.81	3.98

Table 8

Analysis Of Variance Of Perceived Competency Social On Rewards and Standards (N=96)

Source	df	MS	F	p
Main Effects	3	0.566	0.794	0.501
Rewards	2	0.594	0.833	0.438
Standards	1	0.510	0.716	0.400
Interaction	2	0.510	0.716	0.492
Explained	5	0.544	0.762	0.579
Residual	90	0.713		

Table 9

Mean Scores On Perceived Competency Social Scale

	Standards	No Standards	Total
Performance Contingent Reward	4.00	4.13	4.06
Task Contingent Reward	4.00	3.63	3.81
No Reward	4.13	3.94	4.03
Total	4.04	3.90	3.97

Table 10

Analysis Of Variance Of Perceived Task Importance On Rewards and Standards (N=96)

Source	df	MS	F	p
Main Effects	3	1.250	0.954	0.418
Rewards	2	1.792	1.368	0.260
Standards	1	0.167	0.127	0.722
Interaction	2	1.167	0.891	0.414
Explained	5	1.217	0.929	0.466
Residual	90	1.310		

Table 11

Mean Scores On Scale Of Perceived Task Importance

	Standards	No Standards	Total
Performance Contingent Reward	2.88	2.88	2.88
Task Contingent Reward	2.94	2.69	2.81
No Reward	3.00	3.50	3.25
Total	2.94	3.02	2.98

Table 12

Analysis Of Variance Of Score On Target Task On Rewards and Standards
(N=96)

Source	df	MS	F	p
Main Effects	3	2.177	1.147	0.335
Rewards	2	1.760	0.928	0.399
Standards	1	3.010	1.586	0.211
Interaction	2	2.698	1.422	0.247
Explained	5	2.385	1.257	0.290
Residual	90	1.898		

Table 13

Mean Scores On The Target Task During The Experimental Manipulation

	Standards	No Standards	Total
Performance Contingent Reward	5.06	5.25	5.16
Task Contingent Reward	5.44	5.31	5.38
No Reward	5.13	6.13	5.63
Total	5.21	5.56	5.39

Table 14

Pearson Product-Moment Correlations (N=96)

	Time on Target Task	Perceived Competency Personal	Perceived Competency Social	Perceived Task Importance
Time on Target Task				
r	1.000	0.2365	0.0024	0.2585
p		0.020	0.981	0.011
Perceived Competency Personal				
r		1.0000	0.3791	0.2143
p			0.000	0.036
Perceived Competency Social				
r			1.0000	0.2519
p				0.013
Perceived Task Importance				
r				1.0000
p				

Figure 1

Factorial Design of Study

	Standards	No Standards
Performance Contingent Reward		
Task Contingent Reward		
No Reward		

Appendix A

Attitudinal Scales

PERCEIVED PERSONAL COMPETENCE

How do you think you did when you played this game?

excellent	very good	good	fair	poor

PERCEIVED SOCIAL COMPETENCE

How do you think I would say you played this game?

excellent	very good	good	fair	poor

PERCEIVED TASK IMPORTANCE

How important do you think it is to do well on this game?

Extremely important	very important	important	slightly important	not important

Appendix B

(Original printed on yellow paper)

Appendix C

Appendix D

Differential Instructions

Grouped According to Experimental Condition

		Standards	No Standards
Performance Contingent Reward		If you can find and circle three or more things you will be doing a good job. For a good job I have a special prize for you. (described) You will have enough time. Please stop when the buzzer rings. Then if you have done a good job, I will make out your prize certificate.	For a good job I have a special prize for you. (described) You will have enough time. Please stop when the buzzer rings. Then if you have done a good job, I will make out your prize certificate.
Task Contingent Reward		If you can find and circle three or more things you will be doing a good job. For playing this game I have a special prize for you. (described) You will have enough time. Please stop when the buzzer rings. Then I will make out your prize certificate.	For playing this game I have a special prize for you. (described) You will have enough time. Please stop when the buzzer rings. Then I will make out your prize certificate.
No Reward		If you can find and circle three or more things you will be doing a good job. You will have enough time. Please stop when the buzzer rings.	You will have enough time. Please stop when the buzzer rings.

References

Adams, J. S. (1963). Toward an understanding of inequity. Journal of Abnormal and Social Psychology, 67, 422-436.

Arkes, H. R. (1978). Competence and the maintenance of behavior. Motivation and Emotion, 2, 201-211.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. Psychological Review, 74, 183-200.

Boggiano, A. K., & Ruble, D. N. (1979). Competence and the overjustification effect: A developmental study. Journal of Personality and Social Psychology, 37, 1462-1468.

Calder, B. J., & Staw, B. M. (1975). Self-perception of intrinsic and extrinsic motivation. Journal of Personality and Social Psychology, 31, 599-605.

Davis, R. T., Settlage, P. H., & Harlow, H. F. (1950). Performance of normal and brain-operated monkeys on mechanical puzzles with and without food incentive. Journal of Genetic Psychology, 77, 305-311.

DeCharms, R. (1968). Personal causation: The internal affective determinants of behavior. New York: Academic Press.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology, 18, 105-115.

Deci, E. L. (1972a). The effects of contingent and noncontingent rewards and controls on intrinsic motivation. Organizational Behavior and Human Performance, 8, 217-229.

Deci, E. L. (1972b). Intrinsic Motivation, extrinsic reinforcement, and inequity. Journal of Personality and Social Psychology, 22, 113-120.

Deci, E. L. (1975). Intrinsic Motivation. New York: Plenum Press.

Enzle, M. E. & Ross, J. M. (1978). Increasing and decreasing intrinsic interest with contingent rewards: A test of cognitive evaluation theory. Journal of Experimental Social Psychology, 14, 588-597.

Feingold, G. B., & Mahoney, M. J. (1975). Reinforcement effects on intrinsic interest: Undermining the overjustification hypothesis. Behavior Therapy, 6, 367-377.

Festinger, L. (1957). A theory of cognitive dissonance. Stanford: Stanford University Press.

Greene, D., & Lepper, M. R. (1974). Effects of extrinsic rewards on children's subsequent interest. Child Development, 45, 1141-1145.

Greene, D., Sternberg, B., & Lepper, M. R. (1976). Overjustification in a token economy. Journal of Personality and Social Psychology, 34, 1219-1234.

Hamner, W. C., & Foster, L. W. (1975). Are intrinsic and extrinsic rewards additive: A test of Deci's cognitive evaluation theory of task motivation. Organizational Behavior and Human Performance, 14, 398-415.

Harackiewicz, J. M. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. Journal of Personality and Social Psychology, 37, 1352-1363.

Harlow, H. F., Harlow, M. K., & Meyer, D. R. (1950). Learning motivated by a manipulation drive. Journal of Experimental Psychology, 40, 228-234.

Karniol, R. & Ross, M. (1977). The effects of performance-relevant and performance-irrelevant rewards on children's intrinsic motivation. Child Development, 48, 482-487.

Kruglanski, A. W., Alon, S., & Lewis, T. (1972). Retrospective misattribution and task enjoyment. Journal of Experimental Social Psychology, 8, 493-501.

Kruglanski, A. W., Friedman, I., & Zeevi, G. (1971). The effects of extrinsic incentive on some qualitative aspects of task performance. Journal of Personality, 39, 606-617.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards; A test of the overjustification hypothesis. Journal of Personality and Social Psychology, 28, 129-137.

Lepper, M. R., & Greene, D. (1975). Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. Journal of Personality and Social Psychology, 31, 479-486.

Reiss, S., & Sushinsky, L. W. (1975). Overjustification, competing responses, and the acquisition of intrinsic interest. Journal of Personality and Social Psychology, 31, 1116-1125.

Rosenfield, D., Folger, R., & Adelman, H. F. (1980). When rewards reflect competence: A qualification of the overjustification effect. Journal of Personality and Social Psychology, 39, 368-376.

Smith, M. B. (1968). Competence and socialization. In J. A. Clausen (Ed.), Socialization and Society. Boston: Little, Brown & Co..

Vasta, R., Andrews, D. E., McLaughlin, A. M., Stripe, L. A., & Comfort, C. (1978). Reinforcement effects on intrinsic interest: A classroom analog. Journal of School Psychology, 16, 161-165.

Veroff, J. (1969). Social comparison and the development of achievement motivation. In C. P. Smith (Ed.) Achievement-related motives in children. New York: Russel Sage Foundation.

Weiner, B. (1974). Motivational psychology and educational research. Educational Psychologist, 11, 96-101.

White, R. W. (1959). Motivation reconsidered: The concept of competence. Psychological Review, 66, 297-333.