1979

The effects of systems of reward on subsequent intrinsic motivation.

John Anthony Bates
University of Massachusetts Amherst

THE EFFECTS OF SYSTEMS OF REWARD ON SUBSEQUENT INTRINSIC MOTIVATION

A Thesis Presented

By

JOHN ANTHONY BATES

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

May 1979

Department of Psychology

THE EFFECTS OF SYSTEMS OF REWARD ON

SUBSEQUENT INTRINSIC MOTIVATION

A Thesis Presented

By

JOHN ANTHONY BATES

Approved as to style and content by:

James M. Royer, Chairperson of Committee

Seymour M. Berger, Member

Vonnie C. McLoyd, Member

Bonnie R. Strickland, Department Head
Department of Psychology

ACKNOWLEGEMENTS

I would like to thank the following people for their assistance in the completion of this study.

--Mike Royer, advisor and committee chairperson, whose critical eye and unfailing encouragement forced me to produce something worthy of a graduate degree

--Sy Berger, committee member, whose warm smile and quick wit significantly decreased the anxieties of this often confused and always concerned graduate student

--my wife, Nancy, whose understanding, sacrifice and support allowed my dream of graduate school to become a reality.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The apparent inability of the American educational system to preserve and enhance the interest in exploration and learning that seems to be intrinsic to most children when first entering school has been cited recurrently in the literature of the field (Bruner, 1966; Dewey, 1900; Goodman, 1962). Whether learning for the sake of learning is a phylogenetic characteristic of the human species, the product of continually lengthening schedules of reinforcement, or a high-level skill achieved through the development of some cognitive process, is a debatable, if not empirical, issue. Nevertheless, many long-standing educational procedures have been charged with stultifying learning motivation in children.

The often casual use of a variety of classroom systems of reward to control disruptive behavior or to increase academic achievement has recently been criticized as being one of the chief threats to the desire to learn. For example, Levine and Fasnacht (1974), in citing current research on intrinsic motivation in children, have concluded that one of the most powerful tools available to the teacher for affecting student behavior--the classroom token economy--may lead to nothing more than token learning. Their fear is that the supplying of a student with extrinsic incentives for learning may be an artificial procedure, unlikely to be paralleled outside the classroom, which may ultimately undermine the inherent human desire to learn for the sake of learning.

The evidence against classroom reward systems is not, however, without ambiguities. Some researchers (notably, Feingold & Mahoney, 1975, and Reiss & Sushinsky, 1975) have found no empirical basis for the contention that token economies, at least, harm intrinsic motivation. Further, those researchers who have noted such ill-effects have themselves frequently warned against a premature condemnation of all reward systems (Lepper & Greene, 1976; Lepper, Greene & Nisbett, 1973). Nevertheless, the implicit goal of a responsible educational program must be the production of individuals capable of autonomous learning. Any strategem that increases the learner's dependence on artificial contingencies extant only within the program detracts from that goal. If it has indeed been demonstrated that certain classroom systems of reward often undermine student interest in the learning process, and that these contra-indications are not exclusive to contrived laboratory environments, then a careful re-evaluation of the use of such procedures is clearly mandated.

Before reviewing the methodologies designed to confirm or disconfirm the effects of any system of reward on intrinsic motivation, it is first necessary to establish the parameters of the phenomenon under investigation. The following sections are intended to provide an understanding of the several definitions of intrinsic motivation which have been experimentally operationalized, and to delineate the variety of independent variables that have been demonstrated to affect intrinsic motivation. Finally, an original

experimental investigation of the relationships between rewards and motivation will be discussed. The results of this study will be applied both to an evaluation of several theoretical perspectives on intrinsic motivation, and to the use of reward systems in a variety of classroom contexts.

Definitions of Intrinsic and Extrinsic Motivation

Although infrahuman research may seem to be an inappropriate source for definitions of human intrinsic motivation, it nevertheless provides a useful analog. Berlyne (1966) has suggested two components to inherent curiosity in animals: specific (goal-directed) exploratory behavior, and diversive (novelty-directed) exploratory behavior. Goal-directed behavior in infrahuman research generally describes activities prior to and directed toward receiving extrinsic reinforcers. This description parallels the notion of human extrinsic motivation. That is, humans often engage in activities for the (apparently) sole purpose of receiving tangible rewards.

In contrast, novelty-directed behavior involves the seeking out of "stimulation, regardless of source or content, that offers optimum amounts of novelty, surprisingness, complexity, change or variety" (Berlyne, 1966, p. 26). It has been similarly noted that humans often appear to perform a task for no reason other than the satisfaction inherent to task participation and completion. The process that leads to this behavior in the absence of external rewards has been labeled intrinsic motivation.

Differing theoretical perspectives have modified these simple descriptors of intrinsic and extrinsic motivation to more fully encompass human behavior. Four of these perspectives will next be discussed. It should be noted that, although each position offers a slightly different interpretation of the concepts of intrinsic and extrinsic motivation, none offer clearly unique predictions of their interrelationships with systems of reward.

Theoretical Approaches to Intrinsic and Extrinsic Motivation

Behavioral contrast. The phenomenon of behavioral contrast has been well established in infrahuman research (Freeman, 1971; Reynolds, 1961; Terrace, 1968; and others). Although it is a descriptor of behavior and not a theory, it is included in this discussion, since it offers the most functional approach to intrinsic motivation.

The general paradigm used to demonstrate contrast involves training an animal to perform a task on dual schedules of reinforcement. Once behavior has stabilized under both schedules, if the frequency of reinforcement in one schedule is changed, the response rate under the other schedule will change in the opposite direction. Negative contrast is said to have occurred when the frequency of reinforcement is increased in one schedule, resulting in a decreased response rate under the other schedule.

Feingold and Mahoney (1975) have argued that experimental demonstrations of the inverse relationship between human intrinsic

motivation and systems of extrinic reward may not be phenomenologically different from the process of behavioral contrast. A detailed critique of their evidence in support of this contention is researved for a following section. However, a brief description of the typical intrinsic motivation research paradigm is necessary for an understanding of their argument.

In general, subjects are requested to perform a task judged to be interesting by the experimenter. For their participation, subjects are then presented with some tangible reward (e.g., money, tokens, symbolic rewards). Either immediately following reinforcement, or at some later time, subjects are permitted to interact with the experimental task in a free-choice mode. The amount of time subjects spend with the task is compared either with time similarly spent by a non-rewarded control group, or with time spent by the subjects themselves in a pre-treatment free-choice period. Rewarded subjects who spend less time during the post-treatment free-choice period than either type of control group are said to have experienced a decrease in intrinsic motivation.

Two issues of import immediately arise from this general research paradigm. First, the concept of intrinsic motivation has been experimentally operationalized as time freely spent on an activity in the absence of extrinsic rewards. Second, such a definition combined with the differential reinforcement procedures outlined above may be interpreted as the human analog of demonstrations of infrahuman behavioral contrast.

Certainly, no guarantee exists in behavioral research that similar behavioral outcomes in different species (rats or pigeons *vs*. humans) are the result of the same underlying process. Nevertheless, the behavioral contrast approach presents perhaps the simplest description of the "ill effects" of extrinsic reward on intrinsic motivation.

Self-Perception. Originally suggested as an operant psychology alternative to Cognitive Dissonance theory (Bem, 1968, 1972), Self-Perception theory posits that motivation is a product of response patterns learned through acculturation. That is, humans learn through experience that other humans engage in activities to a degree dependent on the expected outcomes or payoffs of the activities. These observations lead the individual to the generalization that motives for engaging in any activity may be deduced from the tangible gains to be realized from participation. Thus, if an individual perceives his task participation to eventuate in powerful extrinsic reinforcers, he may view these reinforcers to be the locus of causality for his participatory behavior. Such an individual is said to be extrinsically motivated. If, however, task participation carries with it no (or very little) salient contingent reward, the individual will perceive his participation to be a product of his own volition, and may be said to be intrinsically motivated.

Although it may at first appear logical to assume that rewarding an individual for performing a task he might otherwise have engaged in for intrinsic motives alone should increase the incentive for

continued performance, Self-Perception theory predicts that such rewards will decrease the probability of future task participation if the reward is withdrawn. This will occur because the individual will experience a shift in the perceived locus of causality for his behavior from the original intrinsic motives to the extrinsic reward. Removing the reward will then decrease both motivation and subsequent activity.

Self-Perception theory also carries with it the possibility that such a reward-motivation interaction need not necessarily occur. It will occur to the extent that a reward is perceived to be salient to the task which is to be rewarded. A reward may have other properties--such as conveying a sense of competency--which may overshadow the magnitude of the reward, or the task may be so entertaining that any rewards are perceived as only tangental to participation. It is not inconceivable, for example, that a highly-paid automobile assembly line worker might have a different level of intrinsic motivation than a similarly paid university professor. Given the condition of reward salience, however, Self-Perception theory clearly predicts detrimental effects on intrinsic motivation.

The Overjustification Hypothesis. The overjustification hypothesis is a direct offshoot of Self-Perception theory, and predicts that an individual's intrinsic interest in an activity will be undermined by inducing him to engage in that activity as an explicit means to some extrinsic goal. Oversufficient extrinsic incentives will be perceived as the locus of control for behavior (Lepper, Greene &

Nisbett, 1973). According to this approach, the nature of an extrinsic reward is a variable with little or no effect on intrinsic motivation. Oversufficiency of reward, then, describes the condition of receiving any reinforcement beyond that which in the past has sufficed to justify task participation, even if the "reinforcement" is insubstantial or merely symbolic.

This perspective of intrinsic motivation differs slightly from the above interpretation of Bem's Self-Perception theory. Lepper, et al., have defined intrinsic motivation as the process leading to task participation in the absence of perceived, salient, unambiguous and sufficient extrinsic rewards. While Bem's approach may be interpreted to suggest that an individual's re-evaluation of locus of causality is, at least in part, a function of task and reward parameters, Lepper, et al., have argued that <u>any</u> extrinsic contingency beyond that which is currently maintaining <u>any</u> behavior will interact with intrinsic motivation. This prediction, combined with corroborative research, has led to a criticism of contractual classroom techniques which offer rewards irrespective of differing initial levels of individual interests in activities.

<u>Personal Causation</u>. DeCharms (1968) has suggested that intrinsic motivation may be distinguished from extrinsic motivation by relative feelings of personal causation. A person who derives satisfaction from having accomplished something through individual effort will perceive himself to be the <u>origin</u> of his own behavior. This perception of personal causation will foster a high level of intrinsic

motivation. Alternatively, a person who primarily dervies satisfaction from the possession of objective rewards which result from his efforts will perceive himself to be a pawn under the control of those extrinsic contingencies. This perception will foster high levels of extrinsic motivation.

Personal Causation makes a prediction similar to those of the previously mentioned theoretical perspectives in regard to the interaction of intrinsic and extrinsic factors. That is, intrinsic motivation will decrease if an extrinsic reward is obtained in a situation where the individual normally perceives himself to be the origin of his behavior. However, Personal Causation makes the additional prediction that intrinsic motivation will increase if expected rewards for task completion are witheld. This will occur, according to DeCharms, due to a necessary re-evaluation of the origins of the behavior. This second prediction seems logically untenable, since its application would suggest that the aforementioned autobmobile assembly line worker would find his job more intrinisically motivating if his pay-envelope were discovered to be empty! This probably absurdity aside, Personal Causation seems to differ from the other theoretical perspectives only to the extent that it provides labels (pawn v. origin) for the perceptual outcomes of reward-motivation interactions.

The four perspectives presented above do not exhaust the alternative interpretations of intrinsic motivation. They reflect a continuum of increasing reliance on hypothetical constructs and

cognitive processes to describe and explain the same behavioral outcomes. In addition, they provide a general framework within which the results of research on intrinsic motivation may be evaluated.

The following sections will briefly review much of this research according to the independent variables which have been demonstrated to have the greatest affect on intrinsic motivation. Before doing so, however, it is necessary to clearly differentiate between two of the most often manipulated variables: contingent and noncontingent rewards. In all cases where a reward or reinforcer has been described as contingent, this term refers to a schedule of reinforcement based on a subject's level of performance on the experimental task. A noncontingent reward is one which has been delivered based only on the subject's participation in the experimental task. The importance of this distinction will become apparent when the implications of the various methodologies for classrrom systems of reward are discussed.

Variables Affecting Intrinsic Motivation

Contingent expected rewards. An early study by Harlow (1950) provides some insight into the interaction between human intrinsic and extrinsic factors, if again from an infrahuman perspective. Harlow presented rhesus monkeys with a latch-puzzle that could be opened only by following several prescribed steps. The monkeys showed great interest in the puzzle, and quickly learned to solve it. Several of the monkeys were then deprived of food for 22 hours. The

experimenter baited the puzzle-latch with a bit of food in the presence of the deprived animals and returned it to them. The monkeys reportedly attacked the previously mastered puzzle without regard to the manipulations necessary to open it. When it was opened and the food eaten, the monkeys demonstrated no interest in continued play with the puzzle.

DeCharms (1968) has cited this study as evidence of the ill effects of extrinsic reinforcers, and as support for a theory of cognitive re-evaluation of the locus of causality for behavior. However, alternative hypotheses for the behavior of the rewarded monkeys include fatigue, satiation, and an increase in arousal (due to deprivation) far above the optimum level, which hindered rather than enhanced activity (Berlyne, 1966; Yerkes & Dodson, 1908). Moreover, to ascribe the cognitive capacity to evaluate complex motives for behavior to monkeys (however intelligent) seems more anthropomorphic than empirical. Still, this infrahuman research closely parallels the results of other studies involving the effects of contingent expected rewards on human behavior.

Deci (1971) has rejected the infrahuman evidence of an interaction between intrinsic and extrinsic factors on much the same basis as described above. He has therefore attemped to demonstrate such an interaction effect within the framework of Self-Perception theory. In his first series of experiments, Deci (1971) hypothesized that different rewards may have different effects according to the interpretation of the individual. Specifically, money may be

interpreted as a causal agent for an activity originally highly intrinsically motivating.

To test his hypothesis, Deci presented 24 college students (who were satisfying course requirements) with a puzzle-solving task (Soma - a commerical puzzle composed of seven differently shaped pieces which may be fitted together to form a nearly infinite variety of configurations). Subjects were asked to reproduce several configurations which had been drawn on paper for them. Twelve subjects were instructed that they would receive one dollar for each correctly solved puzzle. The other twelve were instructed only to solve as many puzzles as they could, with no mention of a reward.

After the experimental session, all subjects were given an 8-minute free-choice period, wherein they could continue solving puzzles or read a variety of magazines. The control (unpaid) subjects spent significantly more of their free-choice time on puzzle-solving than the experimental (paid) subjects. This, according to Deci, indicated a decrease in intrinsic motivation due to a shift in the perceived locus of causality for task participation.

Deci then attempted to replicate his results in a field setting. His subjects were eight college students who worked as headline writers for a bi-weekly newspaper. Half the subjects were paid $ .50 per headline written during the experimental phase, and half were paid nothing. After termination of the payment, unpaid workers wrote significantly more headlines than paid workers. This was again reported as evidence that contingent expected rewards decrease

intrinsic motivation.

Although the Deci (1971) research suggests that "something happened" as a result of experimental manipulation of contingencies, it is difficult to ascribe that "something" exclusively to a change in intrinsic motivation. Deci assumed that his puzzle-solving task was intrinsically motivating for all subjects. His only attempt to verify that assumption was a post hoc measure of students' attitudes toward the task. All students indicated that the task was enjoyable. However, this result may also be explained in terms of Dissonance theory. That is, subjects who were not rewarded rated the task as highly as the rewarded subjects because they needed to justify their participation. Further, if one accepts the notion that extrinsic rewards decrease intrinsic motivation, one should expect to find subjects who were rewarded to rate the task as less enjoyable than did the nonrewarded subjects. This contradiction may reflect only a low correlation between the attitudes subjects express to experimenters in pencil-and-paper surveys and their observed behaviors, or it may reflect deeper theoretical problems. In addition, it is not unreasonable to doubt the level of intrinsic motivation involved in Deci's field study. Headline writing seems less an inherently interesting exercise than it does a tedious, boring task.

Another study conducted by Deci (1972a) was intended to demonstrate no conceptual discrepancies between the Inequity theory of Adams and Self-Perception theory. Adams (1963) has suggested that the observed level of task performance is a function of the degree

of inequity in the input/outcome ratio perceived by an indivdual. His research (Adams, 1963; Adams & Rosenbaum, 1962) demonstrated that a person paid by the hour will increase his performance if the pay appears to be inequitably large, and that a person paid piecework will decrease his performance if the pay appears to be inequitably large. Thus, where intrinsic motivation is defined as level of task performance, it can be manipulated according to the schedule of reinforcement and the intensity of the reward.

Deci presented 96 subjects with the same Soma puzzle used previously. Each subject solved puzzles in one of six conditions: a) not rewarded, b) rewarded with money before the free-choice period, c) rewarded with money after the free-choice period, or, d), e) and f), verbally rewarded in combination with one of the first three. In all money-rewarded contitions, subjects were aware of the forthcoming reward.

The predictions were essentially identical to those of Deci's 1971 studies, with the exception that subjects who were paid before the free-choice period were expected to spend a greater proportion of free-choice time on puzzle-solving as a result of a perceived inequitable input/outcome ratio. That is, these subjects would percieve that they were paid more money than the task deserved, and would, consequently, attempt to compensate for the overpayment by a high level of puzzle-solving activity during the free period.

The results supported Deci's predictions. The unpaid subjects spent a significantly greater proportion of their free-choice time

on puzzle-solving than either of the paid groups; and, subjects paid prior to the free-choice period spent more time with the puzzles than subjects paid after the free-choice period. (The results of crossing monetary rewards with verbal rewards will be discussed later.)

Although Deci has viewed this research as indicative of no conceptual discrepancies between Inequity and Self-Perception, that conclusion may not be compelling. Self-Perception theory argues that, if a person is given extrinsic rewards for performing a task previously intrinsically motivating, that person may re-evaluate the locus of causality for his behavior, thereby reducing subsequent intrinsic motivation. Inequity theory, in contrast, would predict that rewards are evaluated in relation to expended effort, and that an unreasonably large reward would lead to increased activity (intrinsic motivation). On one hand, Deci's results disconfirm Self-Perception theory, because subjects paid before the free-choice period showed an <u>increase</u> in intrinsic motivation. On the other hand, his results disconfirm Inequity theory, since subjects paid after the free-choice period showed a decrease in activity. Deci has offered no theoretical explanation for the implication that <u>expected</u> rewards are more powerful inhibitors of intrinsic motivation than <u>received</u> rewards. How such an explanation might fit neatly into either theory is difficult to imagine. Deci's assumption that the "three to four dollars" (1977a, p. 117) paid to subjects in the rewards-before condition constituted an unreasonable renumeration is probably also unwarranted (especially in these days of rampant

inflation). It seems the only justification for this was a post hoc evaluation of the data. If anything, this study demonstrates the inadequacy of at least two popular theories of intrinsic motivation to deal with all potential task-reward parameters.

Calder and Staw (1975a) have offered an additional criticism of the Deci research. They have pointed out that none of the Deci studies reported performance data for the experimental task. "It is thus unclear whether any change in free time spent on the task is due to a change in intrinsic motivation or merely to differences in performance" (1975a, p. 77). It may be that the contingently rewarded subjects expended more effort on the experimental task, and that their consequent drop in free-choice puzzle-solving time was attributable to satiation or fatigue.

Evidence conflicting with the results of Deci's research has been presented by Reiss and Sushinsky (1975). They have argued that any exposure to a salient rewarding stimulus--such as the promise of monetary reinforcement--prior to experimental treatment, may result in a variety of competing responses. Their "competing response" hypothesis predicts that such perceptual or cognitive distractions will disrupt task performance to the extent that continued participation in a free-choice mode will become less desirable or satisfying.

Reiss and Sushinsky selected six girls and three boys from a kindergarten class to demonstrate that contingent expected rewards (in the context of a more traditional token paradigm) can increase

intrinsic motivation by controlling for competing responses. Subjects won poker chips by listening to one of three target songs over ten reinforcement trials of varying intervals (10 - 100 seconds), and were able to redeem their chips for a variety of toys. Forty-eight hours after training, the subjects were given a 5-minute free-choice period to listen to any of the three songs in an environment very similar to the experimental setting, and in the presence of two adult observers. The results confirmed the hypothesis: subjects listened to their target songs significantly longer than to either of the other two songs.

A crucial problem exists within the methodology presented by Reiss and Sushinsky which makes their interpretation of these results questionable. Post-treatment testing of the subjects' intrinsic motivation to listen to their target songs was conducted in an environment that may have been indistinguishable (for kindergarten children) from that which had previously been associated with potential rewards. Lepper and Greene (1976) have pointed out that this amounts to a discrimination learning paradigm, wherein the subjects had not yet learned to properly discriminate between external contingencies for reinforcement and non-reinforcement. What Reiss and Sushinsky have demonstrated, according to Lepper and Greene, is that the use of token extrinsic reinforcers is effective in maintaining learned behavior provided the external contingencies for potential rewards are maintained.

An additional study of the effects of contingent expected

rewards in the operant tradition deserves mention. Feingold and Mahoney (1975) have cited a number of methodological deficiencies common to previous research in intrinsic motivation: procedures have not been analogous to token economies (experimental sessions have been brief and without mention of a reinforcement effect), there has been a general lack of independent observations, conclusions have often been drawn from marginal results, there has been a neglect of the relevant literature on behavioral contrast, and there has been a lack of control for discrimination learning. For these reasons, Feingold and Mahoney have reported a methodology designed to demonstrate a reinforcement effect, to parallel normal classroom token economies, and to provide continuing measurements sufficient for examining temporal trends and transition states.

Five randomly selected second grade children served as subjects for this study. The experimental task involved connecting dots to form pictures in Follow-the-Dots booklets. An initial baseline measure of the number of dots connected without extrinsic reinforcement was collected for each subject over a 2-week period. Subjects were then rewarded with points, to be exchanged for toys, for connecting dots in excess of their best baseline peformances. Reinforcement continued over four sessions (one week). A second baseline without reinforcement was then established over a 2-week period. Following a 2-week interval of no experimental contact, a third 2-week baseline was recorded. The results indicated an increase in mean responses during the reinforcement procedure,

followed by a drop in responding during subsequent baselines which nevertheless exceeded original baseline measures. Contrary to other research, the children displayed an average increase of 95.08% in performance from Baseline 1 to Baseline 3.

Although Feingold and Mahoney have stated that their results "cast doubt on the assertion that extrinsic reward necessarily undermines intrinsic motivation" (1975, p. 375), a careful look at their data suggests otherwise. Of the five subjects participating, the two who showed the greatest increase from first to third baseline were those who had the lowest level of Baseline 1 activity. Moreover, the subjects who had the highest level of Baseline 1 activity showed the least change on Baseline 3, and demonstrated a trend toward activity decreasing below Baseline 1 if the third baseline period had been extended. If the individual subject performances are considered in this fashion, rather than mean performance across subjects, the outcomes appear well within the predictions of at least the overjustification hypothesis presented by Lepper, et al. (1973, 1975). Specifically, subjects with low initial intrinsic motivation will display an increase in that characteristic following reinforced exposure to a potentially intrinsically motivating activity. Similarly, subjects high in initial intrinsic motivation will tend to decrease activity to the extent that they are confronted with extrinsic rewards contingent on previously intrinsically motivating activity. While correct in their argument that previous research has had methodological flaws, Feingold and Mahoney have not corrected

that methodology to eliminate alternative explanations for their own results.

Clearly, the effects of contingent expected rewards on intrinsic motivation have not been well-established. Methodological inadequacies and conflicts in theoretical interpretations similar to those discussed above may also be noted in the following presentation of research on the effects of noncontingent rewards.

Noncontingent expected rewards. In 1964, Weick offered college students extra course credit for participating in an experimental task. Half of the subjects, before beginning the task, were told they would not receive the expected credits. On a post-task measure of interest, these denied subjects demonstrated a higher degree of task satisfaction than those subjects who had received credits.

DeCharms (1968) has cited this study as evidence of changing levels of intrinsic motivation due to the effects of noncontingent expected rewards. It must be pointed out, however, that Weick interpreted his results from the perspective of Dissonance theory, which would predict that individuals will attempt to justify their participation in an unrewarded activity by re-evaluating upward the desirability of that activity. Kruglanski, Friedman and Zeevi (1971) have approached this problem of conflicting interpretations (which is not unique to noncontingent rewards), and have argued in favor of Self-Perception theory as the most parsimonious explanation of the relationship between noncontingent task inducement and subsequent task enjoyment.

In their 1971 research, Kruglanski, et al., hypothesized a higher quality of task performance and motivation in the absence of noncontingent incentives. All subjects were volunteers between the ages of 15 and 16. They were presented with five tasks which included Zeigarnik measures, suggesting titles for a literary paragraph, composing a story from a list of vocabulary words, answering questions about a newspaper story, and recalling a list of nonsense syllables. Half of the subjects were told they would receive a reward (a tour of the Psychology Department of the Tel-Aviv University) contingent only on task participation, and half were not. Those subjects who did not receive the noncontingent incentive spent significantly more time on-task, and expressed greater task interest and satisfaction than those who were promised the reward.

That the experimental tasks used in this study were hardly of the variety normally considered intrinsically motivating is especially important to those concerned with the efficacy of classroom token economies. Since tokens are generally given to improve low baseline performance (in tasks not likely to be engaged in independently), it may seem reasonable to question on the basis of these results whether extrinsic incentives may be the best means of eliciting a desire behavior. However, the possibility remains that the effects reported in this study were less a function of a re-evaluation of the locus of causality, and more a function of the distractive qualities of the expected reward. In addition, although the results indicated a qualitative inferiority of responses for the

rewarded subjects, the presence of extrinsic noncontingent incentives also resulted in "a tendancy to peform the task in the shortest, fastest, most parsimonious way possible" (1971, p. 615).

Finally, it is important to note that token economies do not generally reward participation in a task as an exclusive response topography. While some incentive may be necessary to initiate a behavior, continued reinforcement is usually made contingent upon reaching established performance criteria. The Kruglanski, et al., research is representative of the literature on noncontingent rewards in that the only criterion for reinforcement was participation in the experimental task. This is, of course, the defining feature of noncontingent rewards, and is not posed as a methodological inadequacy. The fact that this distinction does exist between that body of research and the normal token paradigm, however, severly limits the applicability of its results to the classroom token economy, although perhaps not to other classroom systems of reward.

Lepper, Greene and Nisbett (1973) performed a field study with preschool children to test the effects of noncontingent rewards from the perspective of the overjustification hypothesis. Children who demonstrated high intrinsic interest in a drawing activity during baseline observations in their clasrooms were selected as subjects for the experiment. These subjects were blocked by degree of initial interest in the activity and assigned randomly to one of three treatment conditions: expected reward, unexpected reward, and no

reward. Rewards consisted of a gold star and red ribbon attached to a card labeled "Good Player Award."

The experimental task involved asking the children already observed to have a high interest in drawing to draw pictures (individually and in private) for the experimenter. A post-treatment measure of normal classroom drawing activity demonstrated decreased interest for those subjects who had received a noncontingent expected reward in the experimental session. Subjects in the other two conditions demonstrated no change in their level of activity. In addition, the experimenters noted a qualitative inferiority for the expected-reward subjects' pictures as opposed to the pictures of the other groups.

Lepper and Greene (1975) replicated the earlier Lepper, et al., research with preschool children. In addition to the expected-unexpected reward conditions, this study included surveillance-nonsurveillance via television monitor as independent variables. Subjects were chose in a method similar to the 1973 study. The experimental task was puzzle-solving, and the rewards were free play with one of several attractive toys. The results paralleled those of the earlier study: subjects in the expected reward condition showed less interest in puzzle-solving after treatment than subjects in the unexpected reward condition; and, subjects under surveillance showed less interest than subjects not under surveillance. These results were judged to be confirmation of the overjustification hypothesis.

Lepper, et al., (1973) were careful to point out that their data do not "suggest that contracting to engage in an activity will always or even usually result in a decrement in intrinsic interest in the activity" (p. 136). Althouth others (Levine & Fasnacht, 1974) have interpreted the Lepper, et al., studies (1973, 1974, 1975) as contra-indicative to token economies, the researhers themselves have made a distinction between the general token economy paradigm and their own methodology. Specifically, token economies are instituted when the level of initial intrinsic interest in an activity is very low, or when the activity is one whose attractiveness becomes apparent only through engaging in it to some minimal level of mastery. In their studies, Lepper, et al., chose subjects who had already demonstrated high interest in the activity on a behavioral measure. Thus, any system of reward may prove to be an unwise strategem for students already possessing high motivation for task participation if reinforcement is dispensed noncontingent with performance.

In the response to criticism that the one-trial reinforcement procedure common to overjustification research is not analogous to the multiple-trial token paradigm (cf., Feingold & Mahoney, 1975; Reiss & Sushinsky, 1975), Greene, Sternberg and Lepper (1976) designed a test of their hypothesis within a more typical token economy environment. Their study took place in an elementary school which emphasized an individualized mathematics program. Normal procedures of this program included a "math lab" and a weekly "Awards Assembly" for dispensing extrinsic rewards, such as

certificates and trophies, to deserving students.

Forty-four fourth- and fifth-grade students were selected as subjects on the basis of the amount of time each had spent within the "math lab" over a thirteen-day period working on four new math activities which had been introduced by the experimenters. Subjects were blocked into eleven groups of four, according to the extent to which they had concentrated on their two most-prefereed activities. From these blocks, subjects were randomly assigned to one of four treatment conditions: 1) differentially reinforced for engaging in either of a subject's two most-preferred activities; 2) differentially reinforced for engaging in either of a subjects' two least-preferred activities; 3) differentially reinforced for engaging in either of two activities selected by each subject on the basis of individual preference; 4) non-differentially reinforced for engaging in any of the four activities.

Reinforcement consisted of one credit toward receiving an award at the "Awards Assembly" for every 3 hours (cumulative) spent on the target tasks during "math lab." The reinforcement procedure continued for 12 days. Withdrawal began with an announcement that credits could no longer be given (because it was thought to be unfair to other students), but that the children were still encouraged to use the activities during their lab sessions. Withdrawal continued for 13 days.

In order to demonstrate a reinforcement effect, baseline, treatment and withdrawal phases were compared within each of the three

differentially reinforced groups. In all cases, time spent on target tasks increased during treatment and decreased during withdrawal. The mean time spent on-task during withdrawal was then compared with the time spent during baseline, within each group. Further, these withdrawal data were compared with time spent by non-differentially reinforced subjects on tasks matched according to level of preference expressed during baseline. For example, the time spent by subjects in the high-interest experimental group during withdrawal was compared with the time spent by non-differentially reinforced subjects on tasks for which they had demonstrated the greatest preference during baseline. The time spent by low-interest subjects was compared with time spent by controls on their least preferred tasks. Time spent by choice subjects was compared with time spent by controls on activities with the same rank of preference as established during baseline.

In all three differentially reinforced groups, time spent on-task during withdrawal was less than time spent during baseline (although the difference in the low-interest group was not significant, probably due to a floor effect). Also, except for those who were in the high-interest group, differentially reinforced subjects spent less time on-task during withdrawal than their non-differentially reinforced counterparts. These results were interpreted by Greene, et al., to demonstrate an overjustification effect within the context of a multiple-trial token economy paradigm. They suggested that the inconsistency observed in the high-interest/control

comparison was more a function of methodological than theoretical inadequacy, basing that argument primarily on corroborative findings within a similar paradigm by Colvin (1973).

While the results of this study appear to be strong evidence that a multiple-trial token economy may produce a decrease in intrinsic motivation, it is important to recall the distinction made earlier between contingent and noncontingent rewards. Green, et al., have made continuous reference in this study to the contingent nature of their reinforcement. Nevertheless, the only criterion for dispensing rewards was time spent with the target tasks. Thus, their results offer no clear indication of the phenomenological similarity between decreases in intrinsic motivation due to reinforcement for participation and reinforcement for meeting some performance criterion, assuming that such a similarity indeed exists.

Deci (1972b) has reported an additional study within the same general paradigm as his previous research on contingent rewards, but including several additional variables, and with results conflicting with the early (1973, 1974) Lepper, et al., research. Subjects were confronted with the now-familiar Soma puzzle, and placed in one of six conditions: threatend with punishment for poor performance; given either positive or negative verbal feedback about their performance; rewarded with money contingent either on participation or performance; or, given no reinforcement of any kind. The results suggested that rewards contingent on performance, threats of punishment and negative verbal feedback all decreased subsequent intrinsic

motivation, and that positive verbal feedback increased intrinsic motivation. However, rewards contingent on participation (*i.e.*, noncontingent rewards) had no effect on subsequent behavior.

This study is subject to the same methodological flaws already mentioned for previous Deci studies. Nevertheless, it raises an interesting question concerning the relative effects of contingent and noncontingent rewards. Calder and Staw (1975a) have taken issue with Deci's conclusion that noncontingent rewards do not change intrinsic motivation because subjects are less likely to perceive themselves as motivated by the rewards. Specifically, they have argued that Deci's results merely affirmed the null hypothesis, and that, since one can never know what factor accounts for a lack of change, it is impossible to prove the absence of an effect. They have asked whether "the receipt of noncontingent rewards in this experiment was the same as receiving no treatment at all, or were there other variables which caused the subjects' intrinsic motivation to remain intact" (1975a, p. 78).

Calder and Staw (1975b) have delineated two major problems in accounting for a behavior in terms of intrinsic or extrinsic motivation. First, labeling a behavior as intrinsically motivating begs the question of the nature of the process through which the behavior has become motivating. Second, the methodology currently used leaves open the possibility that alternative explanations may describe the results equally well.

Through their critique of Deci's research (1975a), Calder and

Staw developed an experimental technique to test the interaction of intrinsic and extrinsic motivation by manipulating both intrinsic and extrinsic factors as independent variables, and measuring their effects on task satisfaction and task persistence. They hypothesized an inverse (Self-Perception) effect when a task initially high in intrinsic motivation was paired with a noncontingent monetary reward. That is, subjects would perceive the noncontingent reward to be the locus of causality for their behaviror, and would express dissatisfaction with the task. In contrast, Calder and Staw predicted a direct (reinforcement) effect when the rewarded task was not interesting. In this case, subjects would derive satisfaction from the normally uninteresting task due to the reinforcing quality of the reward.

The experimental task consisted of solving 15 jigsaw-type puzzles. Intrinsic motivation was manipulated by giving half the subjects blank puzzles, while the other half received puzzles that formed interesting pictures. All puzzles were very simple five piece arrangements to minimize the effects of differing puzzle-solving abilities. Half of each group (n=20) of subjects completed the series of puzzles for no pay, while the other half received one dollar on completion of the task. The monetary reward was verbally and visually pointed out to the paid subjects before the experimental session to insure expectancy. After task completion, all subjects were asked to evaluate the task on a 17-point scale ranging from "extremely unenjoyable" to "extremely enjoyable." In addition, as

a behavioral measure, subjects were asked to volunteer for future experiments of a similar nature, but for no reward.

The results indicated a significant interaction between intrinsic and extrinsic factors. Paid subjects given the blank puzzles (extrinsic reward + low intrinsic motivation) rated the task higher and volunteered more often than unpaid subjects given the same puzzles. Conversely, paid subjects given the picture puzzles (extrinsic reward + high intrinsic motivation) rated the task lower and volunteered less often than unpaid subjects given the same puzzles. Moreover, the picture puzzles were actually rated as less enjoyable than the blank puzzles with the introduction of the noncontingent monetary reward.

Although the methodology used to generate these results is not consistent with the token economy paradigm (i.e., single-trial, noncontingent reward with no demonstration of a reinforcement effect), the observed interaction between reward and motivation is of more than theoretical importance. It suggests that a reattribution of the locus of causality for behavior (within the confines of noncontingent reward) is dependent at least in part on the nature of the task for which reward is offered. This relationship may be applied to any number of classroom reward systems to evaluate their potential effects on subsequent behavior.

The results of research investigating the effects of noncontingent rewards seem nearly as contradictory as those reviewed for contingent rewards. However, it is becoming clear that a careful

delineation of task-reward parameters is fundamental to understanding their relationship to intrinsic motivation. Further evidence for this is offered in the following section.

Task-inherent rewards. Kruglanski, et al., (1975b) have investigated another of the conditions for interaction between intrinsic motivation and extrinsic rewards: the "content-consequence" hypothesis. They have suggested that whenever a tangible reinforcer is inherent to a task, its presence should enhance intrinsic motivation. In contrast, if a tangible reinforcer is not normally associated with a task, its introduction may decrease task satisfaction.

To test this hypothesis, Kruglanski, et al., first presented 48 boys (14 to 15 years old) with one of two games. The first was a coin-toss guessing game, the other, a block-building game. Half of the subjects in each group received money contingent on performance in their game, and half received no money. Since winning money was judged to be more commonly associated with the coin-toss game than with the block-building game, Kruglanski, et al., predicted higher intrinsic motivation for the former group of paid subjects than the latter group.

The results supported this prediction. Subjects in the money-intrinsic condition rated their task as more enjoyable and expressed a greater likelihood to re-engage in the task when they were paid than when they were not paid. Subjects in the money-extrinsic condition rated their task higher when no monetary reward was offered. A similar experiment with 15- and 16-year old subjects using

other games (stock market transactions and atheletic games) provided similar results.

In contrast to Deci's interpretation of the interaction between intrinsic and extrinsic factors, Kruglanski, et al., have demonstrated that extrinsic rewards may enhance intrinsic motivation if they are perceived to be inherent to the task content. Unfortunately, no measure was made of the quality of performance for the subjects in these studies. Without such data, we are unsure of the implications for those classroom systems of contingent reward wherein rewards may not be normally regarded as intrinsic to task performance. Should noninherent contingent extrinsic rewards prove to be both quantitatively and qualitatively inferior to inherent contingent rewards, a major re-evaluation of such reward systems would be necessitated. It would follow that any activity could best be motivated (in terms of qualitative performance and resistance to extinction) by creating situations where in participation in the activity could be causally attributed to the activity's content rather than its consequences.

Unexpected rewards. The result of two studies involving unexpected rewards have already been discussed (Lepper, et al., 1973; Lepper & Greene, 1975). It may be recalled that both studies suggested that unexpected rewards had no effect on subsequent intrinsic motivation. Greene and Lepper (1974) have demonstrated a similar lack of effect. However, Kruglanski, Alon and Lewis (1972) offer conflicting results.

In their study, Kruglanski, et al., introduced a series of five

games ("follow-the-leader," "word construction," "song-matching," discover-the-rhyme," and "speed writing") into the activities of four fifth-grade classrooms. Each class was randomly divided into two groups which competed between themselves on the games. No mention was made of any reward for the winning team. At the end of the competition, prizes were given to the members of the winning teams in two randomly selected classrooms. Immediately following the reward, and again one week later, the subjects were asked to evaluate their enjoyment of the activities. Subjects in the prize condition rated the tasks as less enjoyable than subjects in the no-prize condition in both evaluations. Kruglanski, et al., attributed this decline in intrinsic motivation to a perception by the subjects that the competition engaged in was of the sort that normally would yield a reward. The presentation of a reward, even though unexpected, maintained that perception, and the reward was further perceived as the causal agent for behavior.

Kruglanski, et al., have argued that their results may apply to any system which makes salient rewards a normal product of behavior. If rewards come to be expected (as they would, for example, in a token economy), their presence may inhibit intrinsic motivation. In fact, this rationale may explain why the Lepper, et al., studies failed to demonstrate a decrease in intrinsic motivation due to unexpected rewards. The tasks (drawing and puzzle-solving) in which their subjects engaged were not of the sort that normally yield salient rewards. However, too little research in the area of

unexpected rewards has been reviewed to provide a compelling generalization to real-world systems of reward.

Social reinforcers. Proponents of applied behavior analysis have often argued in favor of restricting reinforcers to those which may be more natural to the client's environment, in order to maximize the likelihood of generalization and maintenance of behavior (Ferster, 1971; O'Leary, Drabman & Kass, 1973; and others). Social reinforcers, such as verbal approval, seem to fit the model of natural reinforcers. Three studies have been reviewed which attempt to place the effects of social reinforcers within the schema of intrinsic motivation.

Deci (1971) hypothesized that the effects of social reinforcers (in this case, praise) may not be phenomenologically different from the inherent satisfaction derived from the successful completion of a task, and thus should act to increase intrinsic motivation. To test his hypothesis, Deci replicated the methodology of his original puzzle-solving research, with the exception that praise, rather than money, was the experimental reward. The results indicated a marginally significant difference (.10 level) between praised and unpraised subjects on free-choice time spent with the target task. Deci interpreted these results as supportive of the predicted increase in intrinsic motivation due to verbal reinforcement.

It may be recalled that, in his attempt to demonstrate the compatibility of Self-Perception and Inequity theories, Deci (1972a) again made use of verbal reinforcement, crossed with the presence or absence of monetary rewards. The experimental task was again

puzzle-solving. Although Deci expected a positive effect from praising subjects' performances, only male subjects showed an increase in intrinsic motivation as a function of verbal reinforcement. Female subjects given verbal reinforcement demonstrated no significant change in motivation. Deci hypothesized this to be the case because the experimenter was a male who may have had enough positive interaction with the female subjects before the treatment phase to negate the effects of verbal reinforcement.

Research by Eisenberger (1970) and by Paris and Cairns (1972) lend support to Deci's interpretation of this lack of effect for social approval. Eisenberger demonstrated a deprivation-satiation function for verbal reinforcement in line with these results. Paris and Cairns have suggested that verbal reinforcement is inferior to verbal punishment in promoting learning due to the contextual ambiguity and high frequency of verbal reinforcement in general conversation.

A further study by Deci (1972b) used both positive and negative verbal feedback as consequences of puzzle-solving. It has already been noted that, in this study, positive verbal feedback increased intrinsic motivation, while negative verbal feedback decreased intrinsic motivation. These results, combined with those of the previous two studies, suggest that social reinforcers may be beneficial to intrinsic motivation, provided that their presentation is unambiguously related to task performance.

A Summary of the Evidence

This review of the research investigating intrinsic motivation leads to several tentative conclusions concerning the use of systems of reward within a classroom environment. First, when rewards are made contingent only on participation in an activity (such as dispensing certificates or trophies for mere membership in school-sponsored organizations), this may lead to a decreased interest in those activities, especially if they are, in themselves, entertaining or stimulating enterprizes. If the activities are of the sort which do not encourage a high level of participation (and yet are desirable academic endeavors), such noncontingent rewards may initially enhance student interest.

A second general conclusion is that social reinforcers may contribute to intrinsic motivation if they are salient to the task at hand, and if their presentation is both unambiguous and of a low enough frequency to prevent satiation. These and other more natural reinforcers are probably of greatest value when the task to be rewarded is one not normally associated with a tangible reinforcer.

Generalizations to classroom reward systems based on performance criteria, including token economies, are not clear. In cases where task performance is already closely associated with extrinsic rewards, the absence of appropriate reinforcers would appear to be more damaging to intrinsic motivation then their presence. Whether the inverse may be true for tasks not clearly associated with rewards is a hypothesis which has produced conflicting results (e.g., Deci,

1971, 1972, 1975, vs. Feingold & Mahoney, 1975).

Finally, if, as some research has suggested (Kruglanski, et al., 1972, 1975), repeated pairings of an extrinsic reinforcer with a task leads to a condition wherein task and reward are perceived as inherently inseparable, then one must seriously question the use of such reinforcers in situations where long-term maintenance of behavior is critical. Of course, no reputable behavior analyst prescribes a program where tangible reinforcers are withdrawn with no attempt made to gradually fade in the control of more natural reinforcers. However, this notion of task-inherent rewards may, in part, explain the great difficulty of achieving stable behavior maintenance and generalization that often plagues behavioral researchers (Kopel & Arkowitz, 1975). If so, teachers who routinely dispense rewards for all manner of classroom activities may be well-advised to look elsewhere for methods to enhance the learning motivation of their students.

The methodologies which have engendered these conclusions have often differed widely, and each has suffered from one or more inadequacy. Research which has consistently been interpreted to suggest the harmful effects of token economies on intrinsic motivation has seldom paralleled the token economy paradigm. In fact, few researchers have demonstrated that their rewards were actually reinforcing the target behaviors. Also, aside from Feingold and Mahoney, (1975), Greene, at al. (1976), and Kruglanski, et al. (1972), little effort has been made to investigate the long-term effects of

extrinsic rewards on intrinsic motivation. Even in these cases, "long-term: was limited to one- or two-week investigation.

Methodological inadequacies notwithstanding, an overview of this body of research suggests that, in some cases and for some individuals, extrinsic rewards may be detrimental to intrinsic motivation when that construct is defined as freely choosing to engage in an activity in the absence of those rewards.

The Present Study

The experimental evidence of the effects of systems of reward on intrinsic motivation has generated far more questions than can be investigated in a single study. What, for example, are the antecedents to participating in a task for little or no tangible gain? What specific set of conditions determines whether or not a person will lose interest in an enjoyable activity when a reward is associated with it? Can the construct "intrinsic motivation" be accurately described by a simple behavioral measure of time spent on-task in the absence of reward, or does this description artificially and unnecessarily limit the scope of potential research? Indeed, is it possible that, just as certain tasks may be more entertaining than others, so too may certain individuals tend to be intrinsically or extrinsically motivated, regardless of the task?

From this seemingly endless set of queries, it was determined that at least two are of special and immediate significance to classroom systems of rewards: 1) whether differential reinforcement

interacts with initial levels of interest in an activity, and 2) whether rewarding individuals for participation or performance differentially affects subsequent intrinsic motivation. If it could be demonstrated that contingent and noncontingent rewards interact with the type of activity to be rewarded, this could provide a useful guideline for the dispensation of classroom rewards. Further, if rewarding a student for engaging in or performance on a task were found to play only a small role in determining his future activity with that task, then the issues which have been raised may have little practical value. In contrast, finding such effects would argue strongly for or against the application of any classroom system of reward.

The methodology presented by Calder and Staw (1975) was judged to be especially well-suited for this investigation. It provided a means by which task interest could easily be manipulated, and it was readily adaptable to include both contingent and noncontingent rewards. More pragmatically, this approach allowed the use of a more accessible subject population (college students) than those earlier studies which have concentrated on primary-aged children, and it did not necessitate extensive training of observers, or elaborate controls for experimenter effects.

The methodology, as it has been presented, required some modification. Although Calder and Staw conducted a pretest to demonstrate that picture-puzzles were more interesting than blank-puzzles, they did not demonstate that their monetary reward was in

fact reinforcing behavior. Also, one must question the validity of using potentially reactive pencil-and-paper attitude measures as the dependent variable to determine subject interest and likelihood to re-participate in the target tasks. A well-defined behavioral measure, in addition to these, would provide more compelling evidence of an affect.

Finally, the Calder and Staw approach was not designed to address the issue most critical to future applications of systems of reward in the classroom. Although the results of their study suggested that certain rewards lead to a higher evaluation of certain tasks immediately following presentation of the rewards, Calder and Staw did not investigate the effects of subsequent participation or performance _in the absence of rewards_ on re-evaluation of task interest. A classroom example may best illustrate this issue. Elementary school children are offered some salient reward for time spent practicing the multiplication tables. Since this activity may generally be regarded as low-interest, the reward may enhance task desirability. If the children were to evaluate their interest in the task immediately after receiving their rewards, they would (as a generalization from the Calder and Staw results) rate it higher than if they had not been rewarded. However, this knowledge is of no value in predicting attitudes toward multiplication when rewards are no longer available. Any intrinsic motivation research paradigm that does not supply the information necessary to make this extrapolation fails as a source of evidence for or against systems of reward.

The following study was conducted with these methodological considerations in mind. As in the Calder and Staw experiment, subjects were presented with a series of either picture- or blank-puzzles, and were asked to solve as many puzzles as possible within a specified time period. All subjects rated both how interesting they found each puzzle to be, and how much enjoyment they had derived from the entire task. Subjects in the two puzzle-type conditions received either one nickel for each correctly solved puzzle (contingent reward), a lump sum of nickels dependent only on task participation (noncontingent reward), or no extrinsic reward. One week after this procedure, all subjects were again asked to solve a set of puzzles, this time for no extrinsic reward, to make the same ratings of interest and enjoyment, and to volunteer for future research of a similar nature.

No single theoretical perspective was embraced to predict all the major results of this study. Each perspective reviewed was judged to be either too restrictive to make sense of all the data, or too broad to be empirically testable. Rather, because the between-within groups nature of the design allowed a number of interesting comparisons, the potential individual effects of reward type and interest levels on intrinsic motivation were predicted on the basis of which perspective seemed to offer the most parsimonious explanation. Also, for the purposes of this study, intrinsic motivation was defined as a combination of expressed interest in and enjoyment on a task, task persistence and consequent performance, and

expressed willingness to continue task participation in the clear absence of a salient reward.

Since a demonstration of a reinforcement effect was judged to be critical to the external validity of this research (in the context of applications to classroom procedures), it was expected that significantly more puzzles would be solved, during the first experimental session, by contingently rewarded subjects than by nonrewarded subjects, due to the presence of a contingent monetary reinforcer. No prediction was made for a similar difference between noncontingently rewarded and nonrewarded subjects, since the noncontingently rewarded group was not differentially reinforced for levels of performance. It was thought, however, that if the noncontingently rewarded group performed at a higher level than did the nonrewarded group, this could be interpreted as evidence in support of Adams' Inequity theory (i.e., an oversufficient, temporally-contingent reward yields higher performance).

Four predictions were made for between-groups comparisons of the effects of reward type and task interest on intrinsic motivation during the first experimental session: 1) noncontingently rewarded subjects given picture-puzzles would rate the task lower in interest and in overall enjoyment than would corresponding nonrewarded subjects; 2) noncontingently rewarded subjects given blank-puzzles would rate the task higher in interest and in overall enjoyment than would corresponding nonrewarded subjects; 3) contingently rewarded subjects given picture-puzzles would rate the task higher in interest and in

overall enjoyment than would corresponding nonrewarded subjects; 4) contingently rewarded subjects given blank-puzzles would rate the task lower in interest and in overall enjoyment than would corresponding nonrewarded subjects.

Given a demonstration that solving picture-puzzles was a more interesting task than solving blank-puzzles, the first two predictions parallel the results of the original Calder and Staw research. Their confirmation would re-illustrate the importance of determining task interest before administering rewards for participation. Some explanation is necessary for the other two predictions. Reinforcement contingent on performance on any task serves at least two purposes. First, it acts as a controlling device to maximize the amount of effort expended by an individual engaging in the task. Second, it provides feedback to the individual about how well the task has been mastered, in the form of tangible evidence. Whichever of these properties of reinforcement is judged to be most salient by the individual will determine how the reinforcement affects task interest. In the case of contingently rewarded subjects solving picture-puzzles, it was thought that the relatively interesting nature of the task would enhance the feedback property of the reinforcer. Consequently, interest and enjoyment were expected to increase for these subjects. On the other hand, contingently rewarded subjects solving blank-puzzles were expected to view the reinforcer as the locus of control for their behavior, since their task would be relatively uninteresting, resulting in a decrease in

both task interest and enjoyment.

The more interesting set of results were expected to be generated from the data of the second experimental session. It was difficult, however, to provide clear predictions for more than a few of these outcomes. Contingently rewarded subjects were predicted to complete fewer puzzles in the second experimental session than in the first, whether they had solved picture- or blank-puzzles, since this combination of treatment and withdrawal was essentially an extinction paradigm. Whether this behavior would be paralleled by a decrease in ratings of task interest and enjoyment was not clear. The results of earlier research involving noncontingent rewards (i.e., Greene, Sternberg & Lepper, 1976; Lepper & Greene, 1975; Lepper, Greene & Nisbett, 1973) suggested that the noncontingently rewarded subjects solving picture-puzzles would demonstrate lower interest, enjoyment and performance in the second experimental session than in the first. Too little research has been conducted with low interest tasks to allow predictions to be made for the noncontingently rewarded subjects solving blank-puzzles. It was felt that the same qualities of reward which were predicted to yield high task ratings following the first experimental session could potentially affect performance and ratings in the second session, yielding little difference in outcomes.

The major between-groups comparisons following the second experimental session involved investigating whether withdrawal of contingent and noncontingent rewards interacts with initial task interest to produce differing ratings and performance outcomes

relative to nonrewarded subjects. This set of results is most crucial for an evaluation of the relative merits of contingent and noncontingent rewards. However, no predictions were made for the potentially observed effects, because previous research has offered no single compelling argument from which they may be developed.

The final variable to be investigated was the frequency with which subjects chose to volunteer for a subsequent experiment. Once again, predictions from the literature are contradictory. Still, it was expected that these results would parallel whatever differences were observed in task ratings following the second experimental session, and would thus provide corroborative evidence for the existence of an effect on intrinsic motivation.

The present study departs from the mainsteam of current research in that it was intended to provide a comparison of the effects of both task type and reward procedure on initial and subsequent intrinsic motivation. Moreover, the methodology used in this study has not limited the operationalization of the construct "intrinsic motivation" to any single response parameter, nor was it designed according to the precepts of any single theoretical perspective which might limit the kinds of effects to be investigated. It is altogether likely that one of the reasons for the propensity of conflicting results in previous research is that the researchers have not agreed on a common set of dependent measures. Given this proposition, it was expected that the various measures of intrinsic motivation used in this study would not consistently parallel each other in direction

or degree of change. That this was a potential result should not be interpreted as a weakness in the experimental design. Rather, it should be interpreted as evidence that the decison of whether or not to use some system of reward contiguous with some activity should be based on the desired outcomes of engaging in that activity (*e.g.*, performance *vs*. high interest).

CHAPTER II

METHOD

Experimental Materials

In order to manipulate the variable of task interest, four sets of 24 jigsaw puzzles were constructed. Two of the sets included puzzles of randomly ordered color pictures chosen for their visual appeal from four different sources: Playboy magazine cartoons, photographs of baby animals, photographs of antique automobiles, and prints of unusual art by the painter, Rene Magritte. These pictures were laminated onto pieces of heavy posterboard, and measured 8 X 10 inches. The other two sets contained puzzles that were simply blank pieces of heavy posterboard, also measuring 8 X 10 inches. Both picture-puzzle sets were paired with one of the blank sets, and corresponding puzzles in each of these pairs of puzzle-sets were cut to have pieces of exactly the same shape. All puzzles were limited to five pieces to control for fatigue and differing problem-solving strategies. Each piece of every puzzle was clearly numbered with either a 1, 2, 3, 4 or a 5, and the similarly shaped pieces of the puzzles in each picture-blank pair were numbered in exactly the same fashion.

After all puzzles had been constructed and numbered, a "template" for each puzzle was constructed by tracing the outline of its pieces in their correctly completed positions on a long sheet of brown paper. The outlined shapes of each puzzle-piece were numbered

in exactly the same manner as were the pieces themselves. Hence, "solving" a puzzle required only matching its numbered pieces with the correspondingly numbered shapes on the appropriate template, whether the completed puzzle were a picture or a blank.

Two pencil-and-paper measures were also constructed. The first provided subjects with spaces to indicate the order in which pieces were used to solve each puzzle, and a space in which to rate how interesting the subjects felt each puzzle to be. Ratings of puzzle interest were made on a 7-point Likert-type scale, with 1 indicating a very uninteresting puzzle, and 7, a very interesting puzzle. The second measure required subjects to rate their use of a variety of potential strategies for solving the puzzles, and to rate their overall enjoyment of the task on a 9-point scale, with 1 indicating an extremely unenjoyable task, and 9, an extremely enjoyable task. Only the data accured from this latter enjoyment measure and from the measure of individual puzzle interest were used as indicators of intrinsic motivation for analyses.

A pilot study was conducted prior to the actual experiment, wherein subjects solved all puzzles in each set, to determine if, in fact, the picture-puzzles were more interesting than the blank-puzzles. Only the measure of overall task enjoyment was originally used in this pilot, because it was the primary dependent measure of intrinsic motivation used in the original Calder and Staw study. However, this single measure proved inadequate as an indicator of differing levels of task interest. The requirement of rating the

interest level of each puzzle was therefore added to the experimental procedure to insure that subjects actively considered the potential esthetic properties of each puzzle. The inclusion of this procedure resulted in picture-puzzles being rated as more interesting and enjoyable than blank-puzzles ($p < .05$).

Subjects and Procedure

Initially, 90 college undergraduates enrolled in elementary, educational or adolescent psychology courses at the University of Massachusetts volunteered to participate in what was described as "a study of contextual cues and problem-solving behavior." Of these subjects, 84 completed the first experimental session and returned for the second session. The Time-1 data from the six nonreturning subjects were not included in the analyses. Because three of the nonreturning subjects had been contingently rewarded for their Time-1 performance, and had already been yoked to three noncontingently rewarded Time-1 subjects (see below), the data accumulated from the latter group were also not included in analyses. Therefore, an additional nine subjects were recruited (from the same undergraduate courses), bringing the total number of subjects who completed both experimental sessions to 90 (53 females, 37 males).

Thirty subjects were randomly assigned to each of three conditions: 1) given no reward for participation in or performance on the experimental activity (NR); 2) rewarded with one nickel for each correctly solved puzzle (CR); 3) rewarded with a lump sum of nickels

simply for participating in the experimental activity (NCR). The number of nickels awarded to subjects in the NCR condition was determined by yoking each of these subjects with one of the subjects in the CR condition, without the knowledge of either group of subjects. Whatever amount of money had been earned by a CR subject, according to the number of puzzles he/she had solved, was then awarded to his/her yoked NCR counterpart. Thus, subjects in the CR and NCR conditions received the same mean monetary reward. Within each of these three reward conditions, 15 subjects were randomly assigned to solve either the picture- or the blank-puzzles. Both subjects in each CR-NCR yoked pair solved puzzles from the same puzzle set.

Subjects were run individually over two sessions, 1 week apart. Upon arriving at the experimental room for the first session, all subjects were instructed that the purpose of the study was to determine if different populations use different methods to solve problems. In order to determine this, subjects were told they were to solve a series of simple jigsaw puzzles. The pieces of these puzzles (either picture or blank, dependent on the subject's condition) were contained in envelopes arranged in order beneath the appropriate templates that had been outlined on the brown paper which was secured to a long table in the experimental room. Subjects were instructed to solve the series of puzzles in order, without omitting puzzles or leaving any started puzzles unfinished. Approximately half of the subjects in each Reward-Puzzle condition solved puzzles beginning with #1, through #24, and half began with #24, though #1.

This was done to insure that each puzzle would be solved by at least half the subjects. This order of solving puzzles was reversed in the second session for each subject. Also, the order of exposure to one of the two picture- or blank-puzzle sets was counterbalanced, so that approximately half of the subjects received one set during the first session, and half received the other set.

In addition to solving the puzzles, subjects were told they should record the exact order in which pieces had been used to complete each puzzle. The numbered shapes on the templates were pointed out to the subjects, and they were told that the actual puzzle pieces were numbered in exactly the same way, so that their task would essentially be to match the shapes and numbers. Thus, if piece #4 were the first piece placed on the template for the first puzzle, followed by pieces #1, #5, #2, #3, subjects were told to record the numbers, 4, 1, 5, 2, 3, in that order. Subjects were provided with the aforementioned data sheet on which to record this information. Subjects were next told that, on completion of each puzzle, they should stop, look at the puzzle, and determine how interesting it was to them according to the 7-point scale of interest also provided on the same data sheet. The number from this scale corresponding to their interest level for a given puzzle was to be written in a space on the sheet next to the puzzle number.

Finally, all subjects were told that they would have 4 minutes to solve as many of the puzzles as they could. This time period was determined from the results of the pilot study in which subjects had

been asked to solve all puzzles in each set as quickly as possible. Because subjects in this pilot required, on the average, about 10 minutes to solve all the puzzles in a set, the 4-minute time limit was chosen to control for the possibility of any subject's solving all the puzzles. Subjects were assured that they were not expected to solve all the puzzles, but that they should solve as many as possible within the 4-minute period.

Subjects in the NR condition received no further instructions. Subjects in the CR condition were told that they could earn 5 cents for each correctly solved puzzle, so they should solve puzzles as quickly as possible. Their potential earnings were clearly displayed to them in the form of a plastic cup filled with nickels, which was placed at the end of the table of puzzles opposite to that at which they were to begin. NCR subjects were told they would earn a specified amount of money--dependent on the amount earned by their yoked CR counterparts--simply for participating in the activity, regardless of how many puzzles were completed. This reward was displayed to them in the same manner as for CR subjects. No mention was made as to why the reward was made available.

After giving the proper set of instructions and answering any procedural questions, the experimenter told each subject to begin, and left the experimental room. At the end of the 4-minute period, the experimenter returned and instructed the subject to stop. Once the appropriate reward, if any, had been dispensed, subjects were asked to fill out the questionnaire investigating potential

puzzle-solving strategies, and including the rating of overall task enjoyment. After subjects completed this questionnaire, they were told that the second session would involve exactly the same activity, but with a different set of puzzles. The first session was then terminated.

The second session followed essentially the same procedure as the first session. All rewarded subjects were additionally told only that no money was available to pay them for their efforts in this session, but that they should still try to solve as many puzzles as possible. Subjects who had solved picture-puzzles in the first session were again asked to solve picture-puzzles, using the set of picture-puzzles to which they had not been previously exposed. Blank-puzzle subjects again solved blank-puzzles. The same 4-minute time limit, and the requirements of recording solution orders and puzzle interest were observed in this session that had been in the first. At the end of the puzzle-solving period, subjects were again asked to evaluate their use of problem-solving strategies, and to assess their overall task enjoyment.

Finally, all subjects were provided with a mimeographed statement indicating that the study had thus far been successful, but that more subjects were needed for a third experimental session, involving the same types of materials. However, no compensation of any kind would be available for subjects in this third session. Individuals willing to continue their task participation under this condition were asked to sign their names to a sign-up sheet located outside the

experimental room. After reading this statement, subjects were permitted to leave the experimental room, and the door to the room was closed so that the subjects could not be directly observed by the experimenter. After a reasonable interval, to insure that the experimenter did not interfere with each subject's decision whether or not to volunteer for the third session, the experimenter left the experimental room and noted if that subject had so volunteered. This completed the second experimental session.

CHAPTER III

RESULTS

Since separate predictions and planned comparisons were made for each of the four dependent measures investigated in this study, results of data analyses are grouped according to those measures to which they directly pertain.

Number of Puzzles Completed

Mean numbers of puzzles completed and respective standard deviations for each experimental condition are reported in Table 1. Table 2 contains the results of an omnibus analysis of variance for this measure, across both experimental sessions. Based of the results of this overall analysis, which indicated a significant main effect for Time, and marginally significant Puzzle-Type X Time, Reward X Time and Puzzle-Type X Reward X Time interactions, several planned comparisons were conducted. The primary purpose of these analyses was to demonstrate the existence of a Reward effect within the CR condition during the first experimental session, which would support the prediction that subjects who had received performance-contingent rewards would complete more puzzles than their NR counterparts, due to the reinforcing qualities of their reward. This demonstration was judged to be necessary in order to allow generalizations to be made from the interest and enjoyment data discussed later to classroom systems of contingent reward. A similar comparison was made between NR and NCR first-session performance data to provide

TABLE 1

Performance Data: Puzzles Completed

		First Session		Second Session	
		Mean	s.d.	Mean	s.d.
NR	Pictures	8.87	2.20	10.60	2.20
	Blanks	9.27	1.71	10.60	2.29
CR	Pictures	9.93	1.94	11.87	2.20
	Blanks	8.53	2.10	11.33	1.91
NCR	Pictures	9.60	1.88	10.80	2.21
	Blanks	9.67	2.32	12.00	2.83

TABLE 2

Analysis of Variance: Puzzles Completed

Sources of Variance	Degrees of Freedom	Mean Squares	F Ratios
Total	179		
Between-Subjects	89		
Puzzle-Type	1	.09	<1
Reward	2	8.17	<1
Puzzle-Type X Reward	2	10.27	1.23
Error	84	8.35	
Within-Subjects	90		
Time	1	160.56	153.84**
Puzzle-Type X Time	1	3.19	3.07*
Reward X Time	2	2.77	2.66*
Puzzle-Type X Reward X Time	2	2.52	2.42*
Error	84	1.04	

* $p < .10$

** $p < .001$

potential support for Adams' (1963) notion that (relatively) large rewards based on time-on-task, rather than on the amount of work completed, will lead to increased task effort. Additionally, second-session performance data were compared with first-session data within both the CR and the NCR conditions to investigate the effects of reward withdrawal on the rate of puzzle completion for previously rewarded subjects. Second-session rewarded-subject data were also investigated to determine if reward withdrawal had affected the rate of puzzle completion relative to the performance of NR subjects. Finally, CR and NCR performance data were compared within both experimental sessions to determine whether the modes of reward presentation had differentially affected the numbers of puzzles completed by subjects in those conditions. Except where otherwise noted, these comparisons were conducted using the Bonferroni-*t* method of analysis (Myers, 1972, p. 361), with *EW* = .10 and *t* values converted to *F*'s with the appropriate degrees of freedom.

It was initially predicted that CR-Picture and CR-Blank subjects would complete significantly more puzzles than would their respective NR counterparts, during the first experimental session, due to a reinforcement effect. The result of a planned comparison of first-session performance between the CR- and NR-Picture subjects supported this prediction, $F(1, 84) = 8.10$, $p < .01$; however, as the means reported in Table 1 indicate, CR-Blank subjects completed slightly *fewer* puzzles than did NR-Blank subjects, demonstrating no reinforcement effect.

Although the observed difference in mean performance between NR- and CR-Blank subjects (9.27 vs. 8.53 puzzles completed, respectively) was not statistically significant, the fact that a Reward effect had occurred in one CR group but not in the other suggested that two different processes may have been operating within the CR-Picture and CR-Blank conditions. A post hoc analysis of interaction was therefore conducted on the first-session NR and CR data, using the Scheffe complex-contrast procedure (Myers, 1972, p. 363), with $\alpha = .05$ and the critical $F = 6.30$. The result of this analysis, $Fs = 11.68$, $p < .005$ indicated a significant Puzzle-Type X Reward interaction, suggesting that performance-contingent reward had effected a differential puzzle-completion rate in the two CR conditions. A tenative interpretation of this unexpected result, in combination with the second-session CR-Blank performance data, will be discussed below.

No predictions were made concerning the effects of reward on first-session effort for NCR subjects, although it was suggested that a significant increase in effort for these subjects would be supportive evidence for Adams' Inequity Theory. NCR-Picture and NCR-Blank subjects did complete slightly more puzzles than did NR-Picture and NR-Blank subjects during the first experimental session (9.60 and 9.67 vs. 8.87 and 9.27 mean puzzles completed, respectively). However, planned comparisons of NR- vs. NCR-Picture means, NR- vs. NCR-Blank means, and overall NR vs. NCR means (summing across Pictures and Blanks) demonstrated no significant Reward effects on performance.

Still, these results do not disconfirm the notion of inequity, since they may be alternatively interpreted to indicate that the mean-reward/time-on-task relationship present in this study ($.462/4 minutes) was not judged to be inequitably large by NCR subjects.

In addition to these comparisons of first-session data, several between-session comparisons had been planned to determine how reward-withdrawal had affected puzzle completion rates within the four rewarded conditions. Also, rewarded-subject performance was compared with NR performance within the second experimental session to determine how reward-withdrawal had affected puzzle completion rates relative to the performance of subjects who had not received a first-session reward. Two predictions were made concerning the between-session comparisons of task performance. Based on the confirmation of a reinforcement effect for CR-Picture subjects within Time-1, it was expected that those subjects would complete fewer puzzles in Time-2, since the withdrawal of reward in the second session would function as an extinction paradigm. Because no reinforcement effect was indicated for CR-Blank subjects, the potential effects of reward-withdrawal on this group were not clear. It was also expected that NCR-Picture performance would decrease during Time-2: the bulk of prior research has demonstrated that pairing a contingent reward with a high-interest task may lead to an over-justification effect, manifested by decreased task performance when the reward has been withdrawn. No prediction was made for the effects of reward-withdrawal on performance for NCR-Blank subjects, because

of the scarcity of research evidence involving noncontingent rewards and low-interest tasks.

In contrast to the expectation of decreased task performance following reward-withdrawal in at least two of the four reward conditions, the means reported in Table 1 and the overall *F* statistic for a Time effect reported in Table 2 indicate that *all* experimental conditions completed significantly more puzzles during Time-2 than during Time-1. Although a portion of this increase in performance was probably due to subjects' familiarity with the experimental procedures and a consequent practice effect, the marginally significant interactions also reported in Table-2 indicate that reward-withdrawal actively contributed to Time-2 performance in at least some of the reward conditions. In order to better understand the relationship between reward-withdrawal and task performance, Time-2 rewarded-subject data were compared with the data from their non-rewarded counterparts. No clear predictions were made concerning potential Time-2 performance differences, again because of the lack of compelling evidence in the literature.

The mean numbers of puzzles completed in Time-2 by CR-Picture and CR-Blank subjects were first compared with the average performances of NR-Picture and NR-Blank subjects. CR-Picture subjects completed significantly more puzzles, within Time-2, than did NR-Picture subjects, $F(1, 84) = 11.63$, $p < .005$. This result was somewhat surprising, because, if the performance of the NR-Picture subjects was representative of a normal baseline level for solving

picture-puzzles, then the removal of a performance-contingent reward could reasonably have been expected to lead to performance slightly below that baseline, due to a negative contrast effect. Nevertheless, the absence of a contrast effect for CR-Picture subjects, and their continued high level of performance, is not dissimilar to the result reported by Feingold and Mahoney (1975). In their study, contingently rewarded children also continued to engage in an interesting task at a rate significantly above baseline, even after withdrawal of reward.

CR-Blank subjects also completed more puzzles during Time-2 than did their nonrewarded counterparts (11.33 vs. 10.60 mean puzzles completed, respectively), although a planned comparison of means indicated this difference was not significant. However, the dramatic increase in puzzles completed from Time-1 to Time-2 demonstrated by subjects in the CR-Blank condition (see Table 1), combined with the fact that they had completed fewer puzzles than had NR-Blank subjects during the first experimental session, was thought to provide potenial evidence that reward-withdrawal had actually acted to enhance CR-Blank subject performance. That is, the presence of a performance-contingent reward during Time-1 may have been perceived by CR-Blank subjects to be an unwarranted extrinsic attempt to control their behavior, which acted to depress subject performance. Removal of this reward during Time-2 may then have been analogous to the removal of an aversive stimulus, resulting in a performance rate slightly above baseline due to a positive contrast effect. A post hoc

comparison of first- and second-session performance between the NR- and CR-Blank conditions was conducted using the aforementioned Scheffé complex-contrast procedure ($\underline{\alpha}$ = .05, critical $\underline{F}$ = 6.30) to test this interpretation. The result of this analysis indicated a significant Reward X Time interaction, $\underline{F}s$ = 7.79, $\underline{p}$ < .01, which was judged to support the interpretation that reward-withdrawal had acted to increase performance within the CR-Blank condition. A further discussion of this result as evidence of a positive contrast effect is provided below.

Second-session performance data of NCR-Picture and NCR-Blank subjects were also compared with the data of their NR counterparts to investigate the relative effects of reward-withdrawal on subjects who had received participation-contingent rewards. Based on prior research involving noncontingent reward and high-interest tasks (e.g., Lepper & Greene, 1975; Lepper, Greene & Nisbett, 1975), it was expected that NCR-Picture subjects would complete fewer puzzles than would NR-Picture subjects during Time-2, due to a reattribution of the motive for task participation from the interest qualities of the task itself to the reception of extrinsic reward. No predictions were made concerning the effects of reward-withdrawal on NCR-Blank subject performance.

A planned comparison of the mean numbers of puzzles completed by NR- and NCR-Picture subjects (see Table 1) indicated that the two groups did not significantly differ in second-session performance. This result disconfirmed the predicted effect of reward-withdrawal on

NCR-Picture subject performance, and is in conflict with the fundamental prediction derived from the overjustification hypothesis--that the presence of any salient extrinsic reward beyond that which would normally motivate engaging in an activity will act to decrease the degree of task participation if that reward is withdrawn.

It has previously been noted that NCR-Blank subjects did not differ in performance from NR-Blank subjects during Time-1. A planned comparison of Time-2 performance between these two groups, however, indicated that NCR-Blank subjects completed significantly more puzzles than did their NR counterparts, $\underline{F}(1, 84) = 14.13$, $\underline{p} < .001$. No such effect of reward-withdrawal had been anticipated for NCR-Blank subjects, and whatever reward-motivation relationships may account for this unexpected differential performance increase are not clear.

Finally, planned comparisons of the four reward conditions in both experimental sessions were conducted to determine the relative superiority of one or another of the task-reward combinations for enhancing effort. Within Time-1, noncontingently rewarded subjects significantly outperformed contingently rewarded subjects when engaged in the low-interest (Blank) puzzle-solving task, $\underline{F}(1, 84) = 9.37$, $\underline{p} < .005$, although both groups received the same mean monetary reward ($.427). CR- and NCR-Picture subjects did not differ in performance during Time-1. During Time-2, CR-Picture subjects outperformed their NCR counterparts, $\underline{F}(1, 84) - 8.26$, $\underline{p} < .01$, but NCR-Blank subjects differed only marginally in performance from

CR-Blank subjects (12.00 vs. 11.33 mean puzzles completed, respectively).

In summary, the results of these data analyses indicate that high-interest task performance, as measured by numbers of puzzles completed, was best facilitated by the presence of a performance-contingent reward, both during reward-presentation and after reward-widthdrawal. In contrast, performance on the low-interest task was not enhanced by the immediate presence of a performance-contingent reward, and may have actually been depressed by it. The pairing of a participation-contingent reward with the low-interest task, although it did not lead to significantly superior performance relative to NR subjects during the first experimental session, produced a higher rate of puzzle completion than did the presence of a performance- contingent reward. After reward-withdrawal, this high rate of performance was maintained by NCR-Blank subjects.

Puzzle Interest-Rating

Mean indices of puzzle interest were calculated for each subject and combined in each condition to generate overall interest ratings for both experimental sessions. These group means and respective standard deviations are reported in Table 3. The results of an overall analysis of variance for puzzle interest are included in Table 4. A series of planned comparisons of interest data was conducted, again using the Bonferroni-*t* method of analysis (*EW* = .10), based on the common error terms generated from this overall analysis.

TABLE 3

Puzzle Interest Ratings

		First Session		Second Session	
		Mean	s.d.	Mean	s.d.
NR	Pictures	3.96	.83	3.78	1.05
	Blanks	3.54	.66	3.37	.72
CR	Pictures	3.45	.69	3.61	.90
	Blanks	3.18	.94	2.96	.93
NCR	Pictures	3.50	1.00	3.81	1.06
	Blanks	4.40	1.20	3.69	1.35

TABLE 4

Analysis of Variance: Puzzle Interest Ratings

Sources of Variance	Degrees of Freedom	Mean Squares	F Ratios
Total	179		
Between-Subjects	89		
Puzzle-Type	1	2.58	1.62
Reward	2	3.81	2.40*
Puzzle-Type X Reward	2	1.78	1.12
Error	84	1.59	
Within-Subjects	90		
Time	1	.38	1.36
Puzzle-Type X Time	1	1.12	4.00**
Reward X Time	2	.09	<1
Puzzle-Type X Time	2	.32	1.14
Error	84	.28	

* $p < .10$

** $p < .05$

The first planned comparison involved a test of the replicability of the Calder and Staw (1975) findings that participation-contingent reward may interact with task-type to effect differing levels of task interest. It was predicted that, during Time-1, NR-Picture subjects would rate their task as more interesting than sould NR-Blank subjects, but that NCR-Blank subjects would rate their task as more interesting than would NCR-Picture subjects. Such an interaction between puzzle-type and reward would support the hypothesis that the presence of a participation-contingent reward may evoke a reattribution of the locus of task interest to the reward, rather than to the task, only if the task is normally considered to be of relatively high interest in the absence of reward. Low-interest tasks, in contrast, may be judged to be more interesting when paired with a participation-contingent reward, due to the reinforcing qualities of the reward.

First-session NR interest data were also compared with CR interest data. It was hypothesized that pairing a performance-contingent reward with a high-interest (picture-puzzle) task would enhance task interest, since the reward would be perceived by subjects as being symbolic of their task mastery. CR-Blank subjects, however, were expected to rate their task as less interesting than would NR-Blank subjects, because the relatively uninteresting nature of their task would increase the salience of the extrinsic controlling properties of their reward. It was predicted that this differential perception of reward would be manifested by a significant

first-session Puzzle-Type X Reward interaction of task interest among the NR anc CR conditions.

Second-session interest ratings were compared with first-session data within each of the four reward conditions to determine how reward-withdrawal had affected task interest. Because it was expected that NCR-Picture subjects would attribute their first-session task interest to the reception of reward, it was predicted that reward-withdrawal would further decrease their interest. No clear predictions were made concerning the effects of reward-withdrawal on task interest-ratings for subjects in the remaining three rewarded conditions.

Finally, planned comparisons were conducted between the second-session interest ratings of NR-Picture and NR-Bland subjects and the ratings of subjects in the corresponding reward conditions. No predictions were made concerning these comparisons. However, it has previously been noted that studies investigating the effects of reward on intrinsic motivation should include such information to facilitate evaluations of classroom reward systems. Although a particular task-reward combination may enhance or decrease interest within the environment providing the reward, of greater importance to educators is whether that level of interest is maintained outside that environment, when extrinsic rewards are no longer available, relative to the interest level of individuals who have not received a reward.

Initially, it was predicted that there would be an interaction

between puzzle-type and reward on subject interest-ratings within the NR and NCR conditions, during the first experimental session. Such a result would support the hypothesis that noncontingent rewards may reinforce interest for low-interest tasks, but decrease interest for high-interest tasks due to an overjustification effect. The result of such a comparison of the NR and NCR conditions confirmed this prediction, indicating a significant Puzzle-Type X Reward interaction, $F(1, 84) = 9.91$, $p < .005$. Although nonrewarded subjects rated picture-puzzles as more interesting than blank-puzzles, noncontingently rewarded subjects rated those same blank-puzzles as more interesting than the picture-puzzles (see Table 3).

Interest ratings of the CR-Picture and CR-Blank subjects were also compared with those of the NR-Picture and NR-Blank subjects, respectively, within Time-1. A Puzzle-Type X Reward interaction was predicted, based on the hypothesis that contingent rewards would be perceived differently by subjects in the two CR conditions. That is, the feedback property of the performance-contingent reward, indicating task mastery, was expected to be more salient for CR-Picture subjects, with the result that those subjects would rate their task as more interesting than would NR-Picture subjects. CR-Blank subjects were expected to perceive the manipulative extrinsic property of their reward, which would decrease their task interest below the level of the NR-Blank subjects. However, no significant interaction was observed, disconfirming this hypothesis. In fact, a post hoc application of the Scheffè complex-contrast procedure indicated that

the NR subjects, summing across picture and blank puzzles, rated their tasks as significantly more interesting than did the CR subjects, $\underline{F}s = 10.14$, $\underline{p} < .005$. One interpretation for this result is that reward was perceived as an external control, and acted to decrease task interest, regardless of the task-type.

Data from all four rewarded conditions during the second experimental session were compared with first-session interest ratings to determine the effects of withdrawal of reward on task interest. It was predicted that NCR-Picture subjects would experience a further decrease in interest in Time-2, reflective of their attributing task interest to the opportunity to receive a reward. No clear predictions were made for the NCR-Blank subjects. As the data in Table 3 indicate, NCR-Picture subjects actually found their task to be slightly more interesting after reward-withdrawal. Although this observed increase was not significant, it contradicted the expected effect of reward-withdrawal on interest. NCR-Blank subjects rated their task to be slightly less interesting during Time-2, but this decrease in interest was also nonsignificant. No predictions were made for the same inter-session comparisons within the CR condition. Although mean task interest increased slightly for CR-Picture subjects and decreased slightly for CR-Blank subjects, these differences were not significant.

The results of these comparisons of first- and second-session interest within the CR and NCR conditions may be better evaluated against the Time effects observed within the NR condition. As is

noted in Table 3, both Reward-Blank conditions paralleled the general decrease in interest across sessions observed in the control groups. This trend is understandable, since whatever novelty was perceived by subjects in their first encounter with the experimental materials should have diminished when they returned to engage in a functionally identical task during the second session. However, both Reward-Picture conditions demonstrated an increase in interest over the same interval. This increase is substantiated by the significant Puzzle-Type X Time interaction reported in Table 4.

It may be tentatively suggested that removal of an expected reward had little or no effect on task interest within the low-interest-task groups, relative to the Time-2 interest ratings of the NR-Blank subjects, but did have a positive effect on interest within the high-interest-task groups, relative to the second-session task interest expressed by NR-Picture subjects. This interpretation is corroborated by the planned comparisons of rewarded-condition task interest with the appropriate NR condition within the second experimental session. The mean interest-ratings of both the CR- and NCR-Picture subjects were statistically equivalent to the mean interest expressed by NR-Picture subjects (3.61 and 3.81 vs. 3.78, respectively). However, CR-Blank puzzles were still rated as marginally less interesting than NR-Blank puzzles, $F(1, 84) = 4.50$, $p < .05$. NCR-Blank puzzles, although not rated as significantly more interesting than NR-Blank puzzles, maintained at least the same trend of differential interest than had been observed in Time-1, and were

rated as significantly more interesting, during Time-2, than CR-Blank puzzles, $F(1, 84) = 14.27$, $p < .001$.

The series of planned and post hoc comparisons conducted with first-session interest data supported the hypothesis that participation-contingent rewards may interact with task-type to produce differing levels of task interest, and indicated that performance-contingent rewards may act to decrease task interest regardless of task-type. The results of second-session data analyses, in contrast, were not generally supportive of the contention common in the current literatue that reward-withdrawal further decreases task interest due to a reattribution of motivation from intrinsic to extrinsic factors. Although CR-Blank puzzles were found to be marginally less interesting than NR-Blank puzzles, the remaining comparisons of Time-2 rewarded-condition interest ratings with NR interest ratings demonstrated no apparent long-term detremental reward effects.

Overall Task Enjoyment

As the data in Table 5 indicate, ratings of overall task enjoyment were not generally affected by the presence or absence of reward. The analysis of variance for enjoyment reported in Table 6 demonstrated only a significant main effect for Time, which resulted from tasks being found less enjoyable in the second experimental session than in the first. Since enjoyment ratings had been expected to parallel ratings of task interest, the same planned comparisons that have been reported for the latter set of data were conducted with the

TABLE 5

Task Enjoyment Ratings

		First Session		Second Session	
		Mean	s.d.	Mean	s.d.
NR	Pictures	5.07	1.67	4.80	1.61
	Blanks	5.00	1.20	4.60	1.18
CR	Pictures	5.20	1.52	5.20	1.52
	Blanks	5.33	1.45	4.87	1.92
NCR	Pictures	5.40	1.99	5.47	1.96
	Blanks	5.53	1.60	4.93	1.91

TABLE 6

Analysis of Variance: Task Enjoyment Ratings

Sources of Variance	Degrees of Freedom	Mean Squares	F Ratios
Total	179		
Between-Subjects	89		
Puzzle-Type	1	.94	<1
Reward	2	3.32	<1
Puzzle-Type X Reward	2	.04	<1
Error	84	4.67	
Within-Subjects			
Time	1	3.47	4.51*
Puzzle-Type X Time	1	2.01	2.61
Reward X Time	2	.04	<1
Puzzle-Type X Reward X Time	2	.27	<1
Error	84	.77	

* $p < .05$

former. Surprisingly, none of these comparisons produced significant results. For example, although a first-session Puzzle-Type X Reward interaction for enjoyment within the NR and NCR conditions had been predicted (based on the findings of the 1975 Calder and Staw research and the results of interest-data comparisons in the present study), no such effect was observed. In fact, Picture- and Blank-puzzle tasks in all four rewarded conditions during both experimental sessions were rated as slightly more enjoyable than were the same tasks in the corresponding NR conditions. Nevertheless, all of these differences were found to be nonsignificant.

It is not immediately clear why a measure of task enjoyment was so insensitive to reward effects in this study, especially since a similar measure has been used effectively in previous research, and since ratings of task interest have already been reported to have been significantly affected by the presence or absence of reward. A further consideration of this problem in terms of the most appropriate dependent measures to be used in intrinsic motivation research is included in the discussion of this study.

Volunteering

It was expected that whatever differences in task interest were observed during the second experimental session would be paralleled by differences in rates of volunteering among the six experimental conditions. Because a post hoc comparison indicated that picture-puzzles had been rated across all groups as more interesting than

blank puzzles, $\underline{F}s = 12.43$, $\underline{p} < .001$ (see Table 3), subjects who had solved picture-puzzles were predicted to volunteer more frequently than subjects who had solved blank-puzzles. Also, based on the analyses of task interest data reported above, it was predicted that CR- and NCR-Picture subjects and NCR-Blank subjects would not differ in rates of volunteering from their NR counterparts, and that CR-Blank subjects would volunteer less frequently than would NR-Blank subjects. $\underline{Z}$-tests of proportions (Hays, 1973, p. 305) were conducted to test each of these predictions.

As the proportions of volunteers reported in Table 7 indicate, only within the NR condition did Picture subjects volunteer even marginally more frequently than did Blank subjects, $\underline{z} = 1.54$, $\underline{p} < .07$. Overall, no difference was observed in the proportions of third-session volunteers between Picture and Blank subjects. Investigation of rewarded-condition volunteering using the appropriate NR conditions as standards of comparison revealed several interesting differences. Although NCR-Picture subjects, as had been predicted, did not volunteer at a rate significantly different from NR-Picture subjects, significantly <u>fewer</u> CR-Picture subjects volunteered than did NR-Picture subjects, $\underline{z} = -2.74$, $p < .007$. Also, whereas CR-Blank subjects did not significantly differ from their NR counterparts, significantly <u>more</u> NCR-Blank subjects volunteered than did NR-Blank subjects, $\underline{z} = 2.11$, $\underline{p} < .04$.

Several of these results were in conflict with the predictions based on second-session interest-data analyses. Specifically,

TABLE 7

Numbers (N) and Proportions (P) of Subjects

Who Volunteered for Third Session

		NR	CR	NCR	Overall
Pictures	N	10	5	8	23
	P	.667	.333	.533	.511
Blanks	N	6	8	10	24
	P	.400	.533	.667	.533
Overall	N	16	13	18	47
	P	.533	.433	.600	.522

CR-Picture subjects, although they had rated their tasks to be as interesting as those of the NR-Picture subjects, volunteered significantly less frequently than did their NR counterparts. In contrast, CR-Blank subjects did not differ from NR-Blank subjects in volunteering for a third session, although they had rated their puzzles to be of lower interest. NCR-Blank subjects did not significantly differ from NR-Blank subjects in Time-2 interest ratings, but nevertheless volunteered significantly <u>more</u> frequently. These observed differences in proportions of volunteers appeared to be surprisingly similar to the interest ratings generated from the first experimental session. Post hoc rank-order correlations were calculated between mean task interest and proportions of volunteers for the six groups, for both experimental sessions, to investigate this similarity. Although the correlation between Time-2 task interest and rate of volunteering was not significant, the same correlation using ranked interest from Time-1 <u>was</u> significant, $r_s = .74$, $p < .05$, one-tailed. Because this result was obtained from only six pairs of ranked data, any conclusions based on it must be tentative, at best. Nevertheless, it appears that subjects may have volunteered for an additional session based more on their initial task interests than on their temporally more immediate task experiences.

Additional Post Hoc Comparisons

During the first experimental session, it was observed that rewarded subjects generally exhibited one of two reactions when told of the availability of monetary rewards. Many subjects either responded that payment was not necessary, or indicated that the prospect of receiving a relatively small reward was somewhat humorous. Other subjects expressed pleasure that they would be paid for the experiment, often explaining that the money would buy a cup of coffee or a pack of cigarettes.

Previously, it was argued that reward may be perceived as an extrinsic control of behavior, effecting a decrease in interest, or as a symbol of competency, increasing interest. Based on the observations of subject reactions to the knowledge of a potential reward, it was further hypothesized that not all individuals within each rewarded condition had perceived their rewards in a similar manner. Specifically, rewarded subjects may have perceived their rewards as controlling their behavior--thus increasing output, but decreasing interest--or they may have rejected altogether the controlling aspect of the reward (as well as the artificial control imposed by the experimental environment) and instead performed at a rate lower than nonrewarded subjects, but with little or no effect on their task interest.

Support for this interpretation of subject behavior was sought by calculating the correlations between numbers of puzzles completed and task interest for subjects in each experimental condition. It

was expected that there would be a positive correlation between numbers of puzzles completed and interest in the NR conditions. That is, increased effort in these groups should be reflective of increased interest in the task. However, if the interpretation of the relationship between rewarded-subject interest and effort were sound, it was also expected that effort and interest during Time-1 should be negatively correlated within at least the CR- and NCR-Picture and CR-Blank conditions, and that these correlations would differ significantly from the correlations obtained in the corresponding NR conditions. That is, the more puzzles completed by these subjects, the greater would have been the controlling aspect of the reward, and the lesser the task interest. Whether this relationship would also be demonstrated by NCR-Blank subjects was not clear, since the earlier analyses had indicated that rewards in this group had effected the predicted increase in task interest, and had not resulted in an overall decrease in task effort.

Table 8 includes the observed correlations between interest and effort within the six experimental conditions. As was expected, a moderate positive relationship between these measures was exhibited within the two NR conditions, but the relationship was negative within the CR- and NCR-Picture and CR-Blank conditions. Moreover, based on a Fisher $\underline{r}$-to-$\underline{Z}$ transformation of these data (Hays, 1973, p. 662), it was determined that the correlations obtained from the CR- and NCR-Picture and CR-Blank groups differed significantly from the correlations obtained from their NR counterparts, $\underline{Z} = -2.18$,

$p < .02$; $Z = -3.51$, $p < .001$; and, $Z = -3.74$, $p < .001$, respectively. No significant difference was found between the correlations for NCR- and NR-Blank subjects.

TABLE 8

Correlations between Numbers of Puzzles Completed and Task Interest Ratings: First Session

	Picture-Puzzles	Blank-Puzzles
NR	.365	.474
CR	-.242	-.513
NCR	-.561	.080

Further demonstration of this differential within-group perception of reward was provided by dividing subject interest ratings in each experimental condition into high- and low-performance blocks, based on the number of puzzles solved by each subject during the first session (see Table 9). An overall analysis of variance was then performed on these blocked data--in effect, covarying interest with effort. This analysis demonstrated a significant Reward X Block interaction across both experimental sessions, $F(2, 72) = 4.38$, $p < .025$, as well as first-session Reward X Block interactions between the NR and NCR conditions, $F(1, 72) = 22.28$, $p < .001$, and between the NR and CR conditions, $F(1, 72) = 20.64$, $p < .001$. These results confirmed the expectation that effort would be inversely related to interest when paired with either performance- or participation-contingent reward.

TABLE 9

Mean Puzzle Interest Ratings Blocked

According to First-Session Performance

			First Session Performance Level		Second Session Performance Level	
			High	Low	High	Low
NR	Pictures	Mean	4.41	3.47	4.08	3.35
		s.d.	.60	.84	1.01	1.03
	Blanks	Mean	3.84	3.28	3.44	3.25
		s.d.	.57	.70	.57	.92
CR	Pictures	Mean	3.17	3.53	3.48	3.67
		s.d.	.62	.58	.90	1.00
	Blanks	Mean	2.87	3.58	2.67	3.16
		s.d.	.95	.89	.91	.99
NCR	Pictures	Mean	3.14	4.06	3.27	4.23
		s.d.	1.13	.76	1.25	.64
	Blanks	Mean	4.05	4.30	3.45	4.08
		s.d.	1.04	1.27	1.11	1.62

Since rewards were no longer available in the second experimental session, it was not expected that this differential relationship would continue. However, second-session comparisons indicated the same Reward X Block interactions that had been observed in the first session, with $F(1, 72) = 19.69$, $p < .001$, and $F(1, 72) = 8.00$, $p < .01$, for the NR-NCR and NR-CR comparisons, respectively. These second-session results, when combined with the rank-correlation data reported in the previous section, suggest the existence of a stronger, temporally more resilient reward effect then may seem reasonable for so small a reward. The implications of this effect for classroom systems of reward are discussed in the following section.

CHAPTER IV

DISCUSSION

Each of the theoretical approaches discussed earlier predicted essentially the same general relationship between intrinsic motivation and reward: the presence of a salient extrinsic reward will tend to decrease interest in continued task participation. The results of this study demonstrate little support for so sweeping a prediction. In fact, the presence of an extrinsic reward was found, in some cases, to actually increase either task performance (e.g., CR-Picture subjects), task interest (e.g., NCR-Blank subjects) or the likelihood to continue task participation (also NCR-Blank subjects). Although neither the notion of behavioral contrast, nor Self-Perception theory, nor the overjustification hypothesis, nor Personal Causation Theory has thus received uncompromised support from the reported data analyses, some of these alternative approaches have fared better than the others in providing reasonable explanations for the reward effects that seem directly related to thie particular perspectives. Each of these approaches will next be briefly reviewed, in order of ascending ability to account for the variety of seemingly contradictory results generated from this study.

Personal Causation theory seems to provide the least satisfactory understanding of reward-motivation relationships of any of the four perspectives. This approach, as advanced by deCharms, clearly predicts an increase in intrinsic motivation upon the withdrawal of an expected extrinsic reward, due to a necessary shift in

the perceived locus of causality for behavior. Thus, this position would suggest that rewarded subjects should have had greater task interest during the second experimental session than during the first. Actually, although all rewarded groups increased in effort, expressed task interest decreased somewhat for CR- and NCR-Blank subjects when rewards were removed. Also, of the four rewarded conditions, only the NCR-Blank group volunteered for a third session in a greater proportion than did their control group; CR-Picture subjects volunteered <u>less</u> frequently than controls. This is hardly evidence supportive of the notion that reward-withdrawal makes task participation more intrinsically motivating. Based on these experimental results, Personal Causation theory appears to be an ineffectual source of both predictions about and consistent explanations for interactions between tasks, rewards and motivation.

The potential value of behavioral contrast as a descriptor of the functional relationship between rewards and task performance may be determined by investigating the behavior of those subjects who solved puzzles for performance-contingent rewards. CR-Blank subjects did demonstrate a result surprisingly analogous to the positive contrast phenomenon noted by behavioral researchers when studying the effects of punishment. That is, upon removal of an aversive stimulus, organisms often temporarily respond at a rate slightly above their normal baselines. CR-Blank subjects, in comparison, solved fewer puzzles than did NR-Blank subjects when presented with what was thought to be a performance-contingent reward, and solved

more puzzles than did those same control subjects when the reward was removed. If the positive-contrast analogy is sound, this would suggest that subjects who were offered a relatively small monetary reward contingent on performance on a relatively uninteresting task behaved as if the reward were functionally a punisher, both in the presence of that reward and after it was withdrawn.

Although the notion of contrast provides an interpretation for the unexpected behavior of CR-Blank subjects, it does not offer a convenient explanation for the absence of a contrast effect within the CR-Picture condition. Because the performance of these subjects during the first experimental session indicated a reinforcement effect (i.e., they had solved more puzzles than had NR-Picture subjects), negative contrast should have occurred during the second session, when rewards were no longer available. That is, CR-Picture subjects should then have responded to the altered schedule or reinforcement which decreased reward availability by solving fewer puzzles then were solved by NR-Picture subjects. Instead, CR-Picture subjects continued to perform at a significantly higher rate than their NR counterparts. Thus, while contrast effects were observed in the CR-Blank condition, none were observed in the CR-Picture condition. A strict application of the behavioral contrast paradigm does not provide a consistent explanation for these conflicting results. Because of this lack of consistency, the behavioral contrast approach also appears to be inadequate as a source of predictions concerning the behavioral outcomes of

task-reward manipulations.

Proponents of the third theoretical approach to be discussed, the overjustification hypothesis, argue that the presence of a perceived, salient extrinsic reward beyond that which is currently maintaining behavior will tend to overjustify task participation, with the results that the task will be found to be less interesting, and that participation will become less likely when the reward is no longer available. Given this argument, one might expect that an analysis of Time-1 data in this study would demonstrate lower task interest within the four rewarded conditions than within the appropriate nonrewarded conditions. In fact, although CR-Blank puzzles were rated as being less interesting than NR-Blank puzzles, NCR-Blank puzzles were rated as more interesting than those of the control group. One might also expect, according to the notion of overjustification, that, after reward-withdrawal, less effort would be expended by previously-rewarded subjects than by NR subjects. In fact, NCR-Picture subjects did not differ in performance from NR-Picture subjects, but CR-Picture subjects actually significantly outperformed both these groups, after reward had been withdrawn. What these results suggest is that overjustification-like effects may or may not occur, dependent to a large extent on the mode of reward presentation.

The Puzzle-Type X Reward interaction on interest found by Calder and Staw (1975) and replicated in this study with NR and NCR subjects further indicates that more is involved in determining subject

task-interest than the simple presence or absence of a salient extrinsic reward. Clearly, task parameters such as novelty and complexity must also be carefully considered before generalizations may be made concerning how a given reward may affect motivation.

The inherent weakness of the overjustification hypothesis would appear to be its failure to consider either task or reward-presentation parameters as important components of the reward-motivation relationship. Therefore, this approach also fails to provide a sufficiently comprehensive theoretical perspective for predicting reward effects on intrinsic motivation.

The interpretation of Bem's Self-Perception theory provided in the introduction suggests that a given reward may or may not adversely affect motivation, dependent on the degree of interest inherent to a given task. The results of this study seem more supportive of this interpretation than of any of the other three perspectives. Self-Perception does not directly address the issue of the method of presenting a reward; however, the notion that individuals learn to value activities and outcomes based on their personal observations of others in similar situations may be expanded to include the learning of values attributed to reward contingencies without jeopardizing the consistency of this perspective. Specifically, given an equality of task parameters, individuals in this society may have learned that "people who work harder to get more are being 'bribed,' and must not really enjoy their work," and that "people who work harder even though they don't get more for their

efforts must really enjoy what they're doing." This extrapolation from Self-Perception theory seems to best account for the differential effects on all measures of interest of the two modes of reward presentation used in this study. Whether comparisons were made on first- or second-sessions data, subjects who were rewarded only for their participation generally demonstrated more positive indicators of motivation than subjects who were rewarded for their efforts.

This approach also explicates to some degree the unexpected first-session decrease in effort observed for CR-Blank subjects. If those subjects perceived their potential reward as a coercive manipulation to insure their best efforts on an apparently tedious task, many may have reacted by rejecting the reward altogether and performing at a rate that would demonstrate they could not so easily be controlled. This reaction of decreased effort would not have prevented those subjects from receiving some reward; so, despite the absolute effort expended, they should have found, and did find, the task to be less interesting than if they had not been rewarded.

Although this interpretation of Self-Perception theory offers potential explanations for many of the reported results which were not immediately understandable given the other three perspectives, it must be made clear that these explanations evolved out of the necessity to make sense of results that had not been predicted by any of the perspectives. The value of Self-Perception theory, in this study, was not in its ability to act as a source of clear-cut

predictions for the effects of reward and task manipulations on intrinsic motivation. Whether the above post hoc explanations for the observed effects have any merit can only be determined by further research. Rather, the value of this approach, in contrast to the others, appears to be in its flexibility to encompass a wide variety of independent variables that may affect motivation, while maintaining the integrity of its primary assumption--that personal attributions of the motives for engaging in an activity are learned via observations of others. Given the results of this study, it seems clear that any theoretical approach to the notion of intrinsic motivation must exhibit just such a high degree of flexibility to be able to provide a comprehensive understanding of the effects of systems of reward.

Beyond having theoretical implications, the results of this study may be tentatively applied to some classroom systems of reward. Classroom reward systems may be evaluated according to their effects on two desirable behavior outcomes. The first of these outcomes is the immediate increase of performance on some educationally relevant activity; the second, continued engagement in that activity outside the classroom. Subjects in this study who were confronted with a relatively high-interest task (picture-puzzles) performed at a higher rate when rewarded for their performance than when not rewarded, and continued to perform at a high rate within the experimental environment when the reward was withdrawn. Given a classroom task that is potentially interesting but not likely to be engaged in without some

additional incentive, these results seem to indicate that performance-contingent reward may provide a satisfactory method for enhancing performance. However, the results of this study also indicate that this procedure may risk increasing performance within the somewhat restrictive environment of the classroom at the expense both of student interest and of continued voluntary engagement in that activity outside that environment.

The potential dangers that may be associated with certain systems of performance-contingent reward are indicated even more strongly by the measures of interest for CR subjects presented with low-interest (blank-puzzle) tasks. The casual applicaton of performance-contingent rewards to a low-interest but important activity may, in some cases, actually act to decrease both performance and interest, especially if the nature of the reward is such that it may be perceived to be an unjustified, intentional external manipulation.

Of course, the methodology used in this study was only in a very superficial sense similar to the system of performance-contingent reward common to classroom token economies. Treatment (reinforcement) was included in only a single task encounter, and no attempt was made to gradually fade out extrinsic rewards and fade in more natural reinforcers. Formal token systems would doubtlessly include vastly different procedures, and the results of this study should therefore not be generalized to their use. Nevertheless, it is likely that many teachers routinely dispense some form of performance-contingent reward without the benefit of extensive training in

applied behavior analysis, uncognizant of the potential long-term effects of their behavior on the motivational orientations of their students.

The use of participation-contingent rewards, based on the reported results, seems to be less damaging to motivation than the use of performance-contingent rewards. Particularly in the case of low-interest tasks, noncontingent rewards have been demonstrated to increase immediate and long-term interest and effort. When paired with high-interest tasks, such rewards, although not enhancing long-term effort, did not significantly decrease long-term interest or willingness to continue task participation, and <u>did</u> increase immediate task performance.

There exists a fundamental practical problem with participation-contingent rewards, however, which may preclude their wide-spread use in educational settings. Students at all levels in our present educational system are usually evaluated according to the quality of their performance rather than according to their willingness to engage in educational activities. Systems of performance-contingent reward, particularly token economies, have been demonstrated to be effective methods for enhancing the quality of performance. Participation-contingent rewards, in contrast, have been reported in the literature to lead to a decrease in performance quality, even though this study has indicated their advantages as interest enhancers. Therefore, such systems of noncontingent reward seem to be most useful, and least harmful to other educational goals, in situations where

student participation must be initiated by a reward incentive, and where immediate improvement in the quality of participation is not of critical importance.

The variable most often investigated in studies of intrinsic motivation has been the willingness of rewarded subjects to engage in an activity when rewards are no longer available. The proportions of subjects in each experimental condition who volunteered for a third, nonrewarded session is the measure most similar to that variable in this study. Although subjects had been expected to volunteer in proportions that would parallel group differences in task interest ratings during the second session, volunteering seemed to be more closely associated with task interest during the first session, when rewards were administered. Further research on the long-term effects of exposure to extrinsic rewards on intrinsic motivation must be conducted to determine if this is a reliable result. Such a finding would be of profound importance to educators: the length of exposure to rewards in this study was brief (subject on-task was 4 minutes, and each experimental session lasted no more than 15 minutes), the magnitude of the reward was small, and one week had elapsed between rewarded and nonrewarded sessions, yet subjects in the CR-Picture and NCR-Blank conditions volunteered in proportions very different from control subjects. If this result is truly indicative of the temporal resilience of reward effects on motivation, it should serve to underscore the care which must be taken when applying a system of reward in a classroom environment. Educational

activities generally require more than 4 minutes of a student's time, and may be continually engaged in for weeks, months or years. The rewards available to teachers for increasing class participation in or performance on such activities may be far more powerful than those used in this experiment. The proper combination of task and reward parameters may have far-reaching impact, positive or negative, on the likelihood that students so rewarded will continue to make full use of their skills when they leave the classroom.

The four dependent measures used in this study did not consistently demonstrate the same reward effects. This was expected, and supports the argument that contradictory results in the literature may have been the products of researchers investigating different variables, yet labeling each as an indicator of intrinsic motivation. What, then, is a suitable measure of intrinsic motivation? The confirmation of the Puzzle-Type X Reward interaction reported by Calder and Staw actually further obscures the direction in which future researchers must pursue this problem. Calder and Staw found their results using overall task enjoyment as the dependent measure. The present study replicated the interaction using a measure of subject interest on each individual puzzle, but demonstrated no main or interaction effects using the same overall task enjoyment measure. Why should asking subjects how much they enjoyed their tasks have generated significant results in one situation and not in another, while asking subjects how interesting their tasks were generated the same significant results?

Certainly, differences in the methodologies of the two studies, however slight, were probably responsible for the contradiction, but this is not an acceptable explication of the basic problem. Why should subjects express enjoyment differently from interest? It makes little intuitive sense to say that a change in methodology made tasks interesting but not enjoyable, or enjoyable but not interesting. Yet, this is exactly what the results of this study seem to indicate. What is a suitable measure of intrinsic motivation? Perhaps the most reasonable course for researchers to take in the future would be to combine as many potential dependent measures as possible into any experimental consideration of intrinsic motivation. If results continue to indicate that slight changes in methodology dramatically influence subjective judgements of task interest, enjoyment, entertainment, etc., and that different measures of the same variable respond differently to the same experimental treatment, then one must question both the practical value and the psychological validity of the construct intrinsic motivation itself. Such results would affirm the thesis tentatively advanced earlier that one must evaluate the potential effects of systems of reward based on the immediate and long-term goals of the activity to be rewarded, and not on an elusive superordinate ideal.

Beyond investigating the existence of intrinsic motivation, future research should be directed toward achieving an understanding of reward-interest-motivation relationships in the clssroom. In a very real sense, the experimental environment should not be equated

with the classroom, since the former describes a one-way experimenter-subject reaction, while the latter involves a complex teacher-student interaction. However, although introducing rewards into a simulated classroom situation may provide greater insight into reward effects, manipulating rewards in an actual classroom may not be advisable. The results both of earlier research and of the present study indicate that certain systems of reward may yield inferior performance and/or decreased interest when paired with certain tasks. Also, there is some evidence that these effects may influence future behavior in the absence of rewards. To replicate these results with children and academically-oriented activities would be ethically unjustifiable.

If, as Greene, et al., have suggested, variations in procedure may in fact determine whether subjects ever think about the reasons for engaging in activities (1976, p. 1231), then future research must delineate these parameters. The notion of dispensing a reward to elicit or enhance a desired behavior pervades the American educational system. Whether that reward exists within a token economy, as Honors Assembly or some more informal classroom procedure, it affects both initial and subsequent behavior. In what direction, to what extent, and for how long a time are the questions which must be answered before the ultimate desirability of any reward system may be finally determined.

REFERENCES

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

Adams, J. S., & Rosenbaum, W. B. The relationship of worker productivity to cognitive dissonance about wage inequities. Journal of Applied Psychology, 1962, 46, 161-164.

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

Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 6). New York: Academic Press, 1972.

Berlyne, D. E. Curiosity and exploration. Science, 1966, 153, 25-33.

Bruner, J. S. Toward a theory of instruction. New York: W. W. Norton and Company, 1968.

Calder, B. J., & Staw, B. M. Interaction of intrinsic and extrinsic motivation: Some methodological notes. Journal of Personality and Social Psychology, 1975, 31, 75-83.

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

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

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

Deci, E. L. Intrinsic motivation, extrinsic reinforcement, and inequity. Journal of Personality and Social Psychology, 1972a, 22, 113-120.

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

Deci, E. L., Cascio, W. F., & Krusell, J. Cognitive evaluation theory and some comments on the Calder and Staw critique. Journal of Personality and Social Psychology, 1975, 31, 81-85.

Dewey, J. The school and society. Chicago: University of Chicago Press, 1900.

Eisenberger, R. Is there a deprivation-satiation function for social approval? Psychological Bulletin, 1970, 74, 255-275.

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

Ferster, C. B. Arbitrary and natural reinforcement. In A. Craziano (Ed.), Behavior therapy with children. Chicago: Aldine-Atherton, 1971, 37-43.

Freeman, B. J. Behavioral contrast: Reinforcement frequency or response suppression? Psychological Bulletin, 1971, 75, 347-356.

Goodman, P. Compulsary mis-education. New York: Random House, 1962.

Greene, D., & Lepper, M. R. How to turn play into work, Psychology Today, 1974, 8, 49-54.

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

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

Kopel, S., & Arkowitz, H. The role of attribution and self-perception of behavior change: Implications for behavior therapy. Genetic Psychological Monographs, 1975, 92, 175-212.

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

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

Kruglanski, A. W., Riter, A., Amitai, A., Margolin, B., Shabtai, L., & Zaksh, D. Can money enhance intrinsic motivation?: A test of the content-consequence hypothesis. Journal of Personality and Social Psychology, 1975, 31, 744-750.

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

Lepper, M. R., & Greene, D. On understanding overjustification: A reply to Reiss and Sushinsky. Journal of Personality and Social Psychology, 1976, 33, 25-35.

Lepper, M. R., Greene, D., & Nisbett, R. E. Undermining children's intrinsic interest with extrinsic rewards. Journal of Personality and Social Psychology, 1973, 28, 129-137.

Levine, F. M., & Fasnacht, G. Token rewards may lead to token learning. American Psychologist, 1974, 29, 816-820.

Myers, J. L. Fundamentals of experimental design. Boston: Allyn and Bacon, Inc., 1972.

O'Leary, K. D., & Drabman, R. Token reinforcement programs in the classroom: A review. Psychological Bulletin, 1971, 75, 379-398.

Paris, S. G., & Cairns, R. B. An experimental and ethological analysis of social reinforcement with retarded children. Child Development, 1972, 43, 717-729.

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

Reynolds, G. S. Behavioral contrast. Journal of the Experimental Analysis of Behavior, 1961, 4, 57-71.

Terrace, H. S. Discrimination learning, the peak shift, and behavioral contrast. Journal of the Experimental Analysis of Behavior, 1968, 11, 727-741.

Weick, K. E. Reduction of cognitive dissonance through task enhancement and effort expenditure. Journal of Abnormal and Social Psychology, 1964, 68, 533-539.

Yerkes, R. M., & Dodson, J. D. The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative and Neurological Psychology, 1908, 18, 459-482.