In order to clarify the conditions under which material rewards have a detrimental effect upon children's later interest in the rewarded task, the effect of a reward for drawing was measured with 24 preschool children ranging in age from 34 to 40 months. The children were grouped as high or low in initial interest on the basis of observation of time spent drawing, and half the children in each group were given an expected reward for drawing, while the other half received no reward. Time spent drawing and "quality" of drawing were measured one week and seven weeks later. The high initial interest children who received a reward lost interest when observed a week later, while the low interest rewarded children gained interest. By seven weeks both groups returned to their original levels. At the time of the reward, high interest rewarded subjects drew more drawings but of poorer quality than did the unrewarded high interest children. Low interest children who were rewarded also drew more than their unrewarded counterparts, but quality was not affected. (Author/BD)
The Effect of External Reward on Interest and Quality of Task Performance in Children of High and Low Intrinsic Motivation

Kathryn Kernodle-Loveland and J. Gregory Olley
University of Massachusetts at Amherst

This paper is based upon a masters thesis conducted by the first author under the direction of the second author. Partial support for the research was provided by a faculty research grant from the University of Massachusetts to the second author. An earlier version of this paper was presented at the biennial meeting of the Society for Research in Child Development in New Orleans 1977. The authors are grateful to Jackie Ledger, Laurie Teraspulsky, and Gene Risi for their help in the conduct of the study and to Ann Russell and the staff of the Human Development Laboratory School for their many hours of help and cooperation. Morton Harmatz and Howard Gadlin provided many helpful suggestions throughout the study. Requests for reprints should be sent to Gregory Olley, Department of Psychology, Tobin Hall, University of Massachusetts, Amherst, Massachusetts, 01003.

Running head: Effect of External Reward on Task Performance
Effect of External Reward

Abstract

In order to clarify the conditions under which material rewards have a detrimental effect upon children's later interest in the rewarded task, the effect of a reward for drawing was measured with 24 preschool children. The children were grouped as high or low in initial interest on the basis of observation of time spent drawing, and half the children in each group were given an expected reward for drawing, while the other half received no reward. Time spent drawing and "quality" of drawing were measured, one week and seven weeks later. The high initial interest children who received a reward lost interest when observed a week later, while the low interest rewarded children gained interest. By seven weeks both groups returned to their original levels. At the time of the reward, high interest rewarded subjects drew more drawings but of poorer quality than did the unrewarded high interest children. Low interest children who were rewarded also drew more than their unrewarded counterparts, but quality was not affected.
The Effect of External Reward on Interest and Quality of Task Performance in Children of High and Low Intrinsic Motivation

The use of material rewards to enhance learning in a wide variety of subjects has been a standard part of research in learning for many years. However, recently, some specific circumstances have been identified under which material or external rewards have a detrimental effect upon learning. For instance, Harter (1974, 1975a, 1975b) has demonstrated that "mastery motivation" or "the desire to solve problems for the sake of discovering the solution" may be a more effective consequence than social reinforcement in children. Several studies (e.g., McCullers & Martin, 1971; Spence, 1970) have demonstrated that material rewards are less effective than nonmaterial feedback of results in discrimination learning tasks. In probability learning situations, at least one study (McGraw & McCullers, 1974) has shown rewarded children's learning to be inferior to that of children who received feedback only.

Perhaps the investigators who have received the most attention in this area have been Greene and Lepper (1974; Lepper, Greene, & Nisbett, 1973). Their "overjustification hypothesis" states that one's intrinsic motivation for an activity may be undermined by inducing the individual to engage in the activity as an explicit means to an extrinsic goal. If the goal is
Effect of External Reward

Psychologically oversufficient, the individual may come to believe that his or her activities are motivated primarily by the reward, and intrinsic motivation will subsequently decrease. Perhaps because of the apparent similarity of Lepper et al.'s (1973) paradigm to token economies and other explicit behavior change programs, the phenomenon has generated wide controversy (e.g., Feingold & Mahoney, 1975; Levine & Pasnacht, 1974).

Yet most studies of the "overjustification hypothesis" do not provide an adequate analogue for the use of external reinforcement in applied settings. For instance, rewards in applied programs are given contingent upon some specified criterion of performance by the child; whereas, the studies of the overjustification hypothesis have typically involved noncontingent or performance-irrelevant rewards. Karniol and Ross (1977) recently addressed this issue by investigating the effects of contingent vs. noncontingent reward and two levels of feedback about performance on subsequent interest in a "slide game". Children receiving a noncontingent reward who were told they performed well showed less subsequent interest in the task than contingent reward and control children. This result is compatible with the overjustification hypothesis. The children who received a noncontingent reward and feedback of poor performance, however, showed greater subsequent interest than contingent reward and control children. Thus, perceived level of performance and contingency of reward interacted to influence interest.
Another aspect of the Lepper et al. (1973) paradigm which differs from the application of rewards in applied settings is that these authors gave rewards to children who were already very interested in the task. All practical applications of external rewards involve children whose initial intrinsic interest is low, and the rewards are intended to raise performance. To date, published studies have not investigated the effects of the Lepper et al. (1973) paradigm on children of high and low initial intrinsic motivation. Lepper et al. (1973) reanalyzed the data from their initial high interest sample by dividing it at the median of interest. They found an increase in subsequent interest among the lower intrinsically motivated subjects in the unexpected reward condition. No children who were truly low in interest were included, however.

The primary dependent measure used to indicate intrinsic motivation has been the amount of time spent on the task. Only a few studies have involved quality of performance measures. Kruglanski, Friedman, and Zeevi (1971) found rewarded children to score lower than nonrewarded children on tests of creativity. Lepper et al. (1973) presented a drawing activity to preschool age children and found that during the session in which the reward was presented, the expected reward group produced more drawings but drawings of lower quality than did children in the unexpected reward or no reward groups.

In order to examine further the relationship of the
"overjustification hypothesis" to applied uses of external rewards, the present study utilized the Lepper et al. (1973) paradigm and included children high and low in initial intrinsic motivation for drawing. In addition to the usual time-spent-drawing measure of interest, a measure of quality of drawings was included. Since few previous studies have included measures of the effect of one reward session over an extended period of time, this study also included follow-up measures taken one week and seven weeks after the experimental session.

On the basis of the published research, it was predicted that intrinsic interest would diminish for children with initial high interest who received a reward but would increase for children with initial low interest who received a reward. Further, it was predicted that an expected reward would lead to more drawings but drawings of lower quality during the reward session for children with high interest. For children with low initial interest, expected rewards should also lead to many drawings, but with regard to quality little previous research exists, and no specific predictions were made.

Method

Subjects

The participants were 24 children (12 boys and 12 girls) of predominantly white, middle-class backgrounds attending a laboratory preschool at the University of Massachusetts. They ranged in age from 34 to 40 months.
Setting and materials

The recording of the children's activities in the classroom took place in an observation room equipped with one-way mirrors and sound amplification. The 24 children were free to play with the available standard preschool materials without instruction from the eight teachers. The target activity was felt tip pen drawing on 30.48 x 45.12 cm white paper which was available as a choice only at the times that observations were made. The drawing activity was offered as a normal part of the classroom routine.

Procedure

Initial intrinsic motivation was measured by recording the number of seconds that each child engaged in drawing. Initial measures were taken for two and one half hours per morning on five consecutive class days, for a total of 12.5 hours. Children were scored as engaged in the activity whenever they were seated at the drawing table or holding a felt tip pen at the table. Scorer reliability was calculated by correlating the number of seconds recorded by the two observers who were uninformed about the exact nature of the study. Reliability over the four periods of observation ranged from .95 to .97.

Children who spent above the median amount of time drawing (02 seconds) during this period were defined as high in initial intrinsic motivation for drawing, and those below the median were defined as low in initial intrinsic interest. The mean time spent drawing was 1046 seconds for the high interest group and
166 seconds for the low interest group. Half the children (with an equal number of boys and girls) at each level of interest were randomly assigned to the expected external reward condition, and the other half were assigned to the no reward condition.

The experimental sessions that followed were a replication of those described by Lepper et al. (1973). These sessions began three days after the end of the initial observations and lasted two school days. During the experimental sessions, each child was escorted individually to a "surprise room" and offered the opportunity to engage in a drawing activity identical to the one in the classroom. Children in the expected reward condition were promised a "Good Player Award" (a certificate with a ribbon, decorations, and a space for the child's name) for participation. Children in the no reward condition were not offered a prize. Two female experimenters conducted these sessions in the manner described by Lepper et al. (1973) in order to assure that the experimenter who was present while each child drew was unaware of the child's experimental condition. All children were allowed six minutes to draw, and their time spent drawing was recorded by observers who were blind to the experimental condition of the child.

The first follow-up observation in the classroom began five days later, when the drawing activity was reintroduced as a choice. The second follow-up observation began seven weeks after the experimental sessions. Both follow-ups were identical to the
initial observation period in time and procedure. All drawings were kept and scored for form-diversity, one measure of a drawing's "quality." This measure involves a count of the number of different forms present in a drawing (Holman, Goetz, & Baer, 1977). All drawings were scored by two raters and their scores correlated to yield a reliability measure. Over the four measurement periods, reliability in scoring form diversity ranged from .79 to .86.

Summary of Design

The study consisted of a 2 (high vs low interest) x 2 (reward vs no reward) x 3 (initial measure vs first follow-up vs second follow-up) mixed design with two between subjects comparisons and one within subjects comparison. An equal number of boys and girls were in each cell. The dependent measures were the log of the number of seconds spent drawing and the mean form diversity score in each of the three observation periods. In addition, number of drawings and "quality" (form diversity) of drawings during the experimental period were recorded.

Results

The measure of interest in drawing was the number of seconds each child spent drawing. These data were transformed to their logs to produce homogeneous treatment variances and analyzed in a 2 x 2 x 3 analysis of variance. This analysis yielded significant effects of interest ($F = 18.44$, $df = 1/18$, $p < .0009$; of course, children were assigned to groups on the basis of initial interest).
trials, which was a comparison of the initial measure, the one week follow-up and the seven week follow-up ($F = 3.84, df = 2/36, p < .03$), trials X interest ($F = 7.59, df = 2/36, p < .002$), and trials X reward X interest ($F = 21.27, df = 2/36, p < .0009$).

The triple interaction illustrates the finding that the overjustification hypothesis held for the group high in initial interest but not for the group low in initial interest, and by the seven week follow-up, both groups had returned to their initial interest levels. These results are illustrated in Figure 1.

For the high interest subjects, an expected reward led to a significant decrease in interest from the initial measure to the follow-up the week after the reward was received ($p < .05$ using Duncan's multiple range test). By the seven week follow-up, these children's interest had again changed significantly ($p < .05$) by increasing to almost exactly the level of the initial measure. Thus, the overjustification effect was replicated, but the effect did not last until the seven week follow-up. For the children who did not receive a reward, no significant changes occurred over time. At the one week follow-up, the non-rewarded children showed significantly higher interest than the rewarded children ($p < .05$ using Duncan's test).
Effect of External Reward

For the children with low initial interest, the reward had virtually the opposite effect. Rewarded children increased in interest from the initial measure to the one week follow-up \( (p < .05 \text{ using Duncan's test}) \), then decreased at the seven week follow-up \( (p < .05) \). The unrewarded children did not change significantly in interest over the three measures. During the one week follow-up, the difference in interest between the rewarded and non-rewarded children of low initial interest barely missed significance at the .05 level.

The second dependent measure, "quality" of drawings, as measured by the number of different forms used, was subjected to the same 2 x 2 x 3 mixed design analysis of variance. This analysis revealed that reward had no significant effect on "quality", nor did it interact with initial interest level or trials. The only significant effect was a small but consistent tendency for quality to increase over time \( (F = 14.18, \, df = 2/36, \, p < .0009) \).

In order to test the predictions regarding number and quality of drawings during the experimental session, number of drawings (as measured by number of sides of paper used) and "quality" as measured by form-diversity were analyzed for high and low interest children. For the high interest children an expected reward led to significantly more drawings during the experimental session \( (F = 13.85, \, df = 1/11, \, p < .004) \), but the drawings were of significantly lower quality than those of the
Effect of External Reward

children who did not receive a reward ($F = 10.18, \text{df} = 1/11, p < .009$). This finding replicated that of Lepper et al. (1973) for high interest children. For the low interest children, an expected reward also resulted in significantly more drawings during the experimental period ($F = 20.61, \text{df} = 1/11, p < .001$). However, quality of drawing was not affected significantly by reward ($F < 1.0$).

Discussion

The results of this study serve to clarify the circumstances under which rewards have a detrimental effect upon children's subsequent interest. For the children high in initial interest, an expected, noncontingent reward lead to significant reduction in interest one week later. This finding replicates Lepper et al. (1973) and several subsequent studies that have confirmed this phenomenon. For children of low initial interest, the reward served to increase interest a week later, a finding that is very encouraging to those who wish to use material rewards to enhance motivation in applied settings.

Whereas Levine and Fasnacht (1974) used the Lepper et al. (1973) findings to discourage the use of token economies, the present findings lend support to the use of material rewards for children who show little initial interest in the task. Of course, the present study is far from a perfect analogue of a token economy. Children in applied behavioral programs receive many contingent rewards over a period of time. The present study...
involved one, noncontingent reward. Nevertheless, the effect of that one reward was still present at least a week later. Any encouragement in the use of rewards must be balanced by the realization that the high interest children's loss of interest also lasted at least a week. Thus, it is essential to take into consideration a child's initial level of interest in a task before applying rewards to groups of children such as school classes.

The more immediate effect of the reward also yielded clarification. During the experimental session high and low interest children who received a reward both drew more drawings than their unrewarded counterparts, but for the high interest children, this increase was at the expense of quality. Since the reward was not contingent upon anything other than participation, no explicit relationship between reward and performance was stated by the experimenter. Of course the demand characteristics of the situation and any past experience with rewards may have influenced the children. It is interesting to note that in the absence of any specific statement by the experimenter, the rewarded children worked toward quantity, rather than quality. One should not conclude from this finding that quality cannot be raised. Fallon and Goetz (1975) and Holman, Goetz, and Baer (1977) gave children opportunities to draw with felt-pens and praised the production of forms that the child had not drawn previously. Using the same scoring system as this study, they found systematic
increases in form diversity. Thus, it is possible to increase
the "quality" of children's drawings, but the reward must be
contingent upon quality.

The results of this study provide further clarification of
the circumstances under which rewards increase or decrease
children's interest. Initial interest level, contingency of the
reward, performance level, and perhaps many other influences affect
later interest. If these factors are considered when rewards are
applied, the likelihood of using rewards to enhance the learning
and intrinsic interest of children will be increased.
References


Effect of External Reward


Figure Caption

Figure 1. Time spent drawing during each of the three observation periods.
INTEREST

HIGH INTEREST - REWARD
HIGH INTEREST - NO REWARD
LOW INTEREST - REWARD
LOW INTEREST - NO REWARD

LOG OF MEAN NUMBER OF SECONDS ENGAGED IN ACTIVITY

INITIAL INTEREST
ONE WEEK FOLLOW-UP
SEVEN WEEK FOLLOW-UP

ONE WEEK
SEVEN WEEK FOLLOW-UP