Attribution Theory has led to predictions that the use of material reward may impair intrinsic motivation in the rewarded activity (decreased play effects). A review of the pertinent literature reveals, however, (a) that attribution research has failed to reliably demonstrate that decreased play effects occur in minimal-trial studies (b) that for what effects are reported, alternative and more plausible explanations are available; and (c) that reinforcement programs (multiple-trial studies) do not produce decreased play effects beyond chance levels. Implications and recommendations based on the review are offered. (Author)
Effects of Reinforcement Programs on Intrinsic Motivation

Leonard W. S. Shinsky
University of Illinois at Chicago Circle

Effects of Reinforcement Programs on Intrinsic Motivation

Leonard W. Sushinsky
University of Illinois at Chicago Circle

In the past several decades, psychologists, educators, and parents have expressed concern that therapeutic and educational programs that employ material rewards may have undesirable effects on the rewarded individuals, such as teaching greed and avarice or impairing the acquisition of intrinsic motivation (cf., O'Leary, Poulos, & Devine, 1972).

Recent formulations of attribution theories have regenerated interest in the possible detrimental effects of extrinsic rewards, especially the detrimental effects of material rewards on intrinsic motivation (e.g., Calder & Staw, 1975; Deci, 1971, 1972; Kruglanski, Alon & Lewis, 1972; Lepper, Greene, & Nisbett, 1973). Although attribution theorists should be credited with having promoted research on this issue, the reaction to this preliminary research is surprising in light of the problematical nature of both the theory and the data (cf. Reiss & Sushinsky, 1975a, b, 1976). The reaction to this research is exemplified by the following enthusiastic statements:

"A first principle is to use extrinsic rewards only when necessary ... we think the burden of proof should be on the person who wants to use extrinsic rewards." (Greene & Lepper, 1974, p. 54).

"Rewarding an intrinsically interesting activity actually makes it less interesting. These findings have important practical implications. ...Although external rewards may motivate the person to perform at the time, they may simultaneously be undermining the person's future interest in performing the activity
for its own sake. This suggests that external rewards should not be used for children who are already motivated and that such rewards should be phased out for the unmotivated child soon after he begins to engage in the activity." (Bem, 1975, p. 540).

"As applied to token economies, the results of these studies indicate the trepidations the practitioner should have before instituting a token program." (Levine & Fasnacht, 1974, p. 817).

"Because of the danger that the use of tokens will decrease the intrinsic satisfaction of activities, they should be avoided unless there is real danger to the person or there is no alternative..." (Levine & Fasnacht, 1974, p. 820).

While some of these commentators apparently have since moderated their stance (e.g., Lepper & Greene, 1976), others have since reiterated their opposition to reinforcement procedures that employ material rewards (e.g., Levine & Fasnacht, 1976). The present paper reviews the research relevant to the issue of the effects of reinforcement programs on intrinsic motivation.

First, however, the following three points should be noted: (1) There are some fundamental ambiguities in the very concept of intrinsic motivation (e.g., see Reiss & Sushinsky, 1975b, 1976). For purposes of this review, however, intrinsic motivation is defined as those measures that have been employed in the pertinent research, which include both self-report measures (e.g., How much do you enjoy the activity? Why do you engage in the activity?) and persistence (skeletal) measures (e.g., observations of time spent on activity). (2) The phenomenon of decreased intrinsic motivation has been termed the "overjustification effect" by some investigators (Lepper et al., 1973). For purposes of this review, however, this latter term is not used since it is associated with both the phenomenon to be explained and a particular explanation. We will employ the term decreased play effect.
as a more descriptive, atheoretical term for any detrimental effects of an environmental stimulus on intrinsic motivation (Reiss & Sushinsky, 1975b). This term encompasses detrimental effects as assessed by both self-report and persistence measures, and was suggested to us by Lepper and Greene's (1975) slogan for characterizing impairment in intrinsic motivation ("Turning play into work"). (3) A reward procedure may be said to have produced a decreased play effect if it is found that after reward training has terminated and the rewards are no longer present, the individual demonstrates less intrinsic motivation in the target activity than do equivalent control subjects who had not been exposed to the reward training. Also, a reward procedure may be said to have produced a decreased play effect if it is found that while an individual is being given rewards in one setting, he concurrently demonstrates less intrinsic motivation in the target activity (compared to appropriate controls) when he is in another setting where rewards are not being given.

Literature Review

Two types of data have been adduced as evidence that intrinsic motivation is impaired by reinforcement programs -- single or minimal-trial reward studies and multiple-trial contingent reinforcement studies. These studies are reviewed separately because, on conceptual grounds, the two types of studies are quite different in at least three ways. First, minimal-trial reward studies do not control for the stimulus novelty effects associated with the early trials of reward (cf. Reiss & Sushinsky, 1975b). In an earlier paper, also responding to concerns over potential detrimental effects of material rewards, O'Leary et al. (1972) reported that their research indicates that the early trials of reinforcement with material rewards may be distracting and are unrepresentative of later trials. Reiss and Sushinsky (1975b, 1976) have recently elaborated theoretically on this finding,
and in their competing response hypothesis have postulated that observations of decreased play effects subsequent to withdrawal of rewards are a consequence of stimulus novelty in the early trials of reward procedures. Furthermore, our research (Reiss & Sushinsky, 1975b; Sushinsky & Reiss, Note 5; Reiss & Sushinsky, Note 2) and research by others (e.g., Davis, Settlage, & Harlow, 1950) supports this temporary distraction effect.

In contrast to minimal-trial reward studies, in reinforcement programs multiple-trials permit the dissipation of stimulus novelty effects.

Secondly, minimal-trial reward studies do not permit an assessment of which behavior is being rewarded. For all we know, compliance to experimenter requests is being rewarded and if we evaluated this behavior during posttests we might find that it does not decrease. (cf. Reiss & Sushinsky, 1976). Reinforcement programs, on the other hand, allow one to observe which specific behavior is strengthened.

Thirdly, minimal-trial reward studies do not permit confidence that for whichever behavior was associated with reward, the reward acted as reinforcer. As is well-known, not all reward procedures are reinforcement procedures; this is one reason the term reinforcement is used in behavior therapy rather than the term reward. Reinforcement programs, of course, permit an assessment of the reinforcing properties of rewards.

**Minimal-trial Studies**

Contrary to what has been supposed, minimal-trial reward studies have not demonstrated reliable detrimental effects of reward on intrinsic motivation. Consider, for example, the first five articles on decreased play effects to appear in the attribution literature; these articles are Deci (1971, 1972), Kruglanski, Friedman, & Zeevi (1971), Kruglanski, Alon, & Lewis (1972), and Lepper, Greene & Nisbett (1973), and among them they reported a total of seven experiments. None of these seven experiments demonstrated a statistically significant detrimental effect on any group of subjects that had received reward prior to posttesting. This fact is surprising
because all seven are widely cited as evidence that rewards impair intrinsic motivation (see, e.g., the discussion sections of the five articles themselves, as well as the following: Bem, 1975; Calder & Staw, 1975; Greene, 1974; Lepper & Greene, 1976; Levine & Fasnacht, 1974; Notz, 1975; Kruglanski, 1975). One of the seven experiments (Lepper et al., 1973) failed to produce meaningful data because subject interdependence during posttests should have precluded the statistical analysis employed to report a significant finding. Two of the experiments found that verbal reward enhances intrinsic motivation (Deci, 1971, 1972), and one experiment found that material reward can enhance intrinsic motivation (Deci, 1972). Two experiments found that promising material reward and then not giving it prior to posttesting had a significant or near significant detrimental effect on intrinsic motivation (Deci, 1972; Kruglanski et al., 1971). One study simply failed to confirm the self-perception hypothesis tested (see Kruglanski et al., 1972, as commented on by Reiss & Sushinsky, 1976). Finally, a pair of studies reported by Deci (1971) question the reliability of the persistence measure of intrinsic motivation because of the failure to replicate the effects of identical control procedures on different groups of college students. From such data, it cannot be concluded that a detrimental effect of reward on intrinsic motivation has been discovered. Nevertheless, this is precisely the conclusion that has been drawn. When reward increased the length of time subjects spent in an activity, the effect was attributed to inferior performance (Deci, 1971) or to feelings of inequity (Deci, 1972) or to the nonsalience of verbal feedback as a reward (Deci, 1972). When reward decreased the length of time subjects spent in an activity, the effect again was defined as inferior performance and attributed to an impairment of intrinsic motivation even though it was not statistically significant (Deci, 1971; Lepper et al., 1973). Findings that reward has no effect on attribu-
tional rating scales have been dismissed (Deci, 1971) even though other findings that reward has an effect on such scales have been interpreted as support for attribution theories (e.g., Calder & Staw, 1975).

Subsequent to the initial articles noted above, a number of single-trial studies did finally report statistically significant detrimental effects of rewards on intrinsic motivation (Calder & Staw, 1975; Kruglanski, 1975; Reiss & Sushinsky, 1975b; Ross, 1975). Even in these subsequent studies, however, the detrimental effect sometimes has occurred at questionable levels of significance (e.g., Staw, Calder, & Hess, Note 4). Moreover, in a recent paper, Reiss (Note 1) has reported three experiments that failed to replicate the previously reported findings that expectation of salient, material rewards impair intrinsic motivation. Since no discernible refinement in training procedures is evident from the initial to the later attribution studies, the question arises as to how many failures to replicate went unpublished against what erroneously appeared to be an initial array of seven replicated findings. Unfortunately, we cannot answer this question.

Multiple-Trial Studies

Three studies have been reported recently as evidence that multiple-trial contingent reinforcement programs are sometimes associated with decreased play effects (Greene, 1974; Johnson, Bolstad, & Lobitz, 1974; Meichenbaum, Bowers, & Ross, 1968).

The Johnson et al. (1974) study is a report on the inappropriate behavior in an untreated setting (either home or classroom) of 19 children who had received treatment for inappropriate behavior in an alternative setting (classroom or home). Decreased play effects would consist of demonstrations of decreased appropriate behavior in the untreated setting concurrent with demonstrations of
increased appropriate behavior in the treated setting. In fact, the data reported in this study fail to confirm the occurrence of decreased play effects because the statistical analyses, performed on 4 separate small samples, typically were either marginal or nonsignificant, control subjects were either unavailable or nonequivalent, regression artifacts were sometimes uncontrolled, and data for some children in the untreated setting were in the direction of enhancement effects. The authors themselves concluded that the data "should be interpreted with caution because of the small sample sizes involved (and) the relatively small magnitude of the effects" and, moreover, the authors indicated that they "are not at all confident that these results will reliably replicate" (p. 14).

The Meichenbaum et al. study reported that material rewards were successfully employed in a reinforcement program to increase appropriate classroom behavior of ten institutionalized delinquent adolescent females. The investigators also reported that when the reinforcement program was only in effect in the afternoons, appropriate classroom behavior in the mornings decreased below baseline level. This finding has been cited as evidence that multiple-trial contingent reinforcement programs that employ material rewards can produce decreased play effects (e.g., Lepper & Greene, 1976).

Actually, in this study there were two separate groups of girls -- one group of four girls (Group I) and one group of six (Group II). Upon inspection of the data presented, it is clear that there is no evidence of decreased play effects for Group II. Aside for the fact that during the baseline period no stable baserate was evident, thus making interpretation difficult in any event, the data points (Meichenbaum et al., 1968, Fig. 2) indicate that during the seven days that treatment occurred in the afternoon, the morning class demonstrated
appropriate behavior at virtually an identical level to that of its own baserate. Meichenbaum et al. also report that on the first occasion that the adolescents were permitted to actually spend the money they had earned for appropriate behavior in the afternoon, appropriate behavior in the morning decreased below baserate levels. This statement is based on a single data point, and this data point actually falls only negligibly below the lowest baseline point, and cannot be distinguished from a chance effect. Furthermore, subsequent treatment days do not demonstrate a decrease from baseline in percentage of appropriate behavior in the morning. Whether this is attributable in part to regression artifacts, to a continuation of unstable operant behavior, or to an enhancement effect is not known. What is known is that the data cannot be considered evidence of decreased play effects, and this interpretation of the data is reasonable if not compelling in the absence of stable baserates.

The data for Group I is also problematical. These data are based on only four girls and are, in any event, uninterpretable. The tendency for the girls to behave more inappropriately in the untreated morning class than during baseline may be no more than a continuation of the consistent trend in this direction that was occurring during the baseline period anyway.

Some authors (Johnson, Bolstad, & Lobitz, 1974) have claimed that statistical analysis of the Meichenbaum et al. data indicates a tendency for appropriate behavior to decrease between the baseline and treatment period for the morning class when all 10 girls are considered. However, this analysis only includes the data from the second week of treatment and thus may be arbitrarily capitalizing on the demonstrated instability of the behavior. Moreover, a more valid comparison would be with an independent group of untreated control subjects, and, of course, under conditions of stable baserates.
Finally, the data from the Meichenbaum et al. study have not been replicated under similar conditions (e.g., O'Leary, Becker, Evans, & Saudargas, 1969) nor under better controlled conditions (Reiss, 1973).

The other pertinent multiple-trial study (Greene, 1974) also fails to demonstrate that contingent reinforcement procedures impair intrinsic motivation. In this study some subjects were reinforced for continuing to play with the two out of four experimental activities they had played most with during baseline assessment (High-Interest group). Other subjects were reinforced for playing with the two activities they had played least with during baseline (Low-Interest group). A third group of subjects were reinforced for playing with any two of the four activities that they indicated they would like to receive rewards for (Choice group). The fourth group, Control group subjects, were reinforced regardless of which activities they played with.

This study cannot be interpreted as having demonstrated decreased play effects for the following reasons:

1. The only statistical analyses (between group comparisons) considered by the author himself to be meaningful in assessing decreased intrinsic motivation, did not demonstrate decreased intrinsic motivation for the High-Interest subjects — who had been expected to demonstrate the greatest impairment in intrinsic motivation.

2. The within-group analysis of a decrease in time on task from baseline to posttest was significant for High-Interest subjects. The author claimed that this decrease is not attributable to satiation effects, and hence may be an indication of a loss in intrinsic motivation. The Control subjects, however, showed a similar decrease from baseline to posttest for time spent playing with high-interest activities. This was the case even though these children had not, as expected, played as much with the high interest activities during treatment as had the High-Interest group.
That is, if the Control group shows natural decreases in time on activity, one should attribute to these same natural processes any similar decreases for experimental subjects — especially a group of subjects that had been playing with the materials for even longer periods of time.

3. The author asserted that if satiation were a reasonable explanation of the decreases in play, then there should be negative correlations for time on task between treatment (or baseline plus treatment) and withdrawal phases. And, in fact, no significant negative correlations were found, this, the author claimed indicating that the satiation explanation was not plausible. The author’s reasoning, however, is confusing, since a satiation explanation does not demand that individuals who played more during treatment (or baseline plus treatment) than did other children should play less than these other children during withdrawal. The satiation explanation merely specifies that the children will play less during withdrawal than they themselves did during treatment (or baseline plus treatment).

4. Further indication that the satiation hypothesis is not merely speculative comes from the Control group’s demonstration of alternation behavior, i.e., over time the high-interest activities were played with less (and the low interest activities more) than during baseline. This same pattern of natural alternation among activities is replicated in the treatment groups when reward contingencies are removed.

5. The yoked control subjects for the Choice group did not replicate the data of the controls for the High-Interest group, thus challenging the reliability of the skeletal measure of intrinsic motivation.

6. The children’s self-reports of their preferences for the experimental activities did not differ as a function of the training they received. Thus, the self-report measures failed to demonstrate that any experimental training condition impaired initial intrinsic motivation.
7. Finally, we note that in this study the children were reinforced for time on activity, not for quality of performance (cf. Reis-Sushinsky, 1975a). What was not assessed was the requisite behaviors a child engaged in in order to persist on the rewarded activity during training. What was the child doing while spending time with the experimental stimulus? As she practicing sloppy play, non-attentive play, frustrated play, anxious play? — all of which could mediate decreased play effects, for example, through Pavlovian conditioning of negative affect, or through the child’s recognition of little enjoyment or poor performance; or which could, if persisting beyond training, be interpreted as evidence of loss of intrinsic motivation (Reiss & Sushinsky, 1976). The question of which behavior topologies were practiced during training becomes pertinent when, as in the Greene (1974) study, a decreased play effect is claimed to have occurred following a training procedure that did not systematically reinforce quality performance. It is not a relevant issue when increased intrinsic motivation (enhancement effect) is found; in that case, the training procedure, by definition, did not reinforce behaviors that successfully competed with intrinsic motivation.

It should be appreciated that we know empirically that it is possible for people who seem to be engaged in functionally equivalent behavior to be learning different response topologies (cf. Logan, 1972). Thus, in cases where decreased play effects occur following time-on-task reinforcement training, the possibility of response topology explanations should be raised. Its post hoc nature is not arbitrary, but simply attests to the fact that the study failed to control for important alternative explanations.

In this context, we note that we do believe that people can come to feel controlled by rewards. We suggest, however, that this is not the consequence of a purely inferential process such as proposed by attribution theory, but rather
a consequence of using rewards to expose individuals to unpleasant activities or
to induce individuals to perform initially pleasant activities in topologically
unpleasant ways (e.g., playing while frustrated or anxious, playing much slower
than usual, or playing more rapidly than usual). Even in these circumstances
the individual could come eventually to feel comfortable with the activity, but
if the activity remains even mildly unpleasant, the individual may attempt to
escape from it if the rewards are not sufficiently powerful and, especially, if
the rewards are removed.

This analysis may explain in part why a professional football player, who
played football "for its own sake" as a child, comes to demand greater and greater
extrinsic rewards in order to continue playing -- and in some cases would rather
sit out the season than play football for a mere $80,000. Playing professional
football is topologically and situationally very different from playing football
with friends. In the former case, the player has to play certain ways, with
certain people, at certain times. Rewarding exposure to potentially unpleasant
behavior such as this might have very different consequences than would reinforcing
the individual for playing football in the manner he had played in childhood; the
former rewards performance under unpleasant conditions but does not necessarily
promote intrinsic motivation (see Reiss & Sushinsky, 1975, Note 3; Reiss &
Sushinsky, 1976). In brief, behavior similarly labelled (e.g., "football playing")
may not be functionally nor topologically equivalent.

As a final comment on multiple-trial studies, we note (a) there are a number
of reinforcement studies that demonstrate increased intrinsic motivation (cf. Reiss
& Sushinsky, 1976); and (b) attribution theorists themselves seem to predict many
more ways that contingent reinforcement will enhance intrinsic motivation than
decrease it (cf. Lepper & Greene, 1976).
Conclusion and Recommendations

The following statements summarize our interpretation of the available data:

1. Decreased play effects are difficult to produce, even in minimal-trial reward studies.

2. The unreliable effects of minimal-trial reward procedures may be related more to distraction/stimulus novelty effects than to attributional processes, and, in any event, do not seem to have relevant implications for reinforcement programs.

3. There is no convincing evidence that reinforcement programs produce decreased play effects beyond chance levels.

From our research, the research of others, and clinical experience, the following two recommendations can be offered:

1. When one wishes to encourage interest in an activity "for its own sake," reinforce the desired topology of the behavior. One implication of this suggestion is that we should not ordinarily reward time on activity since to do so may inadvertently reinforce behaviors (such as sloppy play, hurried play, non-attentive play, or anxious play) which could compete with the acquisition or perpetuation of intrinsic motivation. As we have noted, there are a number of ways that the practice of these types of topologies may lead to subsequent observations (after training) of impaired intrinsic motivation. Furthermore, even if the training led to exactly the topologies reinforced, e.g., sloppy play, observations of these topologies might be interpreted as evidence of decreased intrinsic motivation -- when in fact they are evidence of persistence of reinforced behavior. For example, Wiseman reports that Desmond Morris, the zoologist, once trained an ape to draw and paint, and the animal did so competently. Morris then started rewarding the ape with peanuts for continuing to draw and paint. The ape eventually produced "any old scrawl" to earn peanuts. Although Morris interpreted this behavior as
evidence that the introduction of commercialism had ruined the artist, a more precise interpretation might be that you get what you reinforce—and if Morris had reinforced only quality performance, it is probable that he would have observed quality performance.

A second, related implication of the recommendation to reinforce desired topologies is that one should be cautious in identifying the actual performance that is desired. If one desired to promote intrinsic motivation in schoolwork, then the child should be those behaviors that may lead him or her to enjoy and value educational activities (e.g., performing well, learning about variations in the activities, or what Logan, 1972, terms, developing an expectation of learning something from the activity).

Krumboltz and Krumboltz (1972) report the case of Anne, a fourth-grader who had no difficulty earning A's. After her parents instituted the practice of giving Anne a dime for every A she earned in school, Anne continued to receive A's but her work became careless and "she seemed to be working only for the money and did not seem particularly proud of her work or interested in the material she was learning" (p. 112). Anne's parents, of course, had not meant to reinforce behavior that might be termed "bringing home-papers-with-A's-on them." It may be that the initial trials of the material reward program had some disruptive effects on Anne's performance, or, perhaps the introduction of the program coincided inadvertently with some of Anne's poorer performances. In any event, Anne's parents may have unwittingly reinforced this relatively sloppy behavior by continuing to give rewards to Anne for bringing home "A" papers rather than for quality performances. Fortunately, Anne's mother appreciated what was happening and helped Anne to reinstate her earlier quality behavior by first examining the papers Anne brought home and then reinforcing only the particularly well-done papers.
2. The other recommendation is to use rewards in ways that will not be perceived as punishing, insulting, unfair, or otherwise aversive. Obversely, use rewards in ways that are pleasant and fun. On the one hand, it may appear patronizing to emphasize that rewards should be employed in ways that maximize their reinforcing value. On the other hand, we are aware of a colleague who recently argued that giving his wife money for sex would undermine her interest in sex, thus demonstrating the dangers of material rewards and, it seemed to follow, the likely hazards of reinforcement programs that employ material rewards. Accordingly, it is not gratuitous to suggest the value of associating rewards with pleasant affect. In fact, one of our students is currently analyzing data that appear to demonstrate dramatically the obvious -- although perhaps the not so obvious to our colleague: It is possible to use material rewards in ways that vitiate their reinforcing properties. Specifically, this student has found that children will demonstrate a rapid loss of persistence behavior during a no-reward condition (extinction) if during training they were given material rewards in a manner that defined the rewards as undeserved (Hecimovic, personal communication).

In brief, it might be profitable to heed Bachrach and Quigley's warning of a decade ago: "It is undoubtedly true that the field of behavior therapy, with its worthy social goals, its theoretical simplicity (deceptive though it may be), and its empirical success (under certain circumstances), will attract many psychotechnicians. It is, therefore, a field in danger of being ruined by amateurs" (1967, p. 510).
Reference Notes


References


Reiss, S., & Sushinsky, L.W. Undermining extrinsic interest. *American Psychologist, 1975, 30, 782-783.* (a)
