

## A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation

Edward L. Deci  
University of Rochester

Richard Koestner  
McGill University

Richard M. Ryan  
University of Rochester

A meta-analysis of 128 studies examined the effects of extrinsic rewards on intrinsic motivation. As predicted, engagement-contingent, completion-contingent, and performance-contingent rewards significantly undermined free-choice intrinsic motivation ( $d = -0.40, -0.36$ , and  $-0.28$ , respectively), as did all rewards, all tangible rewards, and all expected rewards. Engagement-contingent and completion-contingent rewards also significantly undermined self-reported interest ( $d = -0.15$ , and  $-0.17$ ), as did all tangible rewards and all expected rewards. Positive feedback enhanced both free-choice behavior ( $d = 0.33$ ) and self-reported interest ( $d = 0.31$ ). Tangible rewards tended to be more detrimental for children than college students, and verbal rewards tended to be less enhancing for children than college students. The authors review 4 previous meta-analyses of this literature and detail how this study's methods, analyses, and results differed from the previous ones.

By 1971, hundreds of studies within the operant tradition (Skinner, 1953) had established that extrinsic rewards can control behavior. When administered closely subsequent to a behavior, rewards were reliably found to increase the likelihood that the behavior would be emitted again, an effect that persisted as long as the reward contingency was operative. When rewards were terminated, the likelihood that the behavior would be emitted eventually returned to the prereward baseline. This general finding led to the widespread advocacy of rewards as a motivational strategy, and behavior-change programs based heavily on the use of rewards were introduced into a variety of applied settings.

In 1971, Deci (1971) argued that some activities provide their own inherent reward, so motivation for these activities is not dependent on external rewards. In line with White (1959), Deci referred to such activities as intrinsically motivated and raised the question of how extrinsic rewards would affect people's intrinsic motivation for these activities. Research reported in that and subsequent papers (Deci, 1971, 1972a, 1972b) revealed that tangible rewards (viz., money) could undermine college students' intrinsic

motivation for an interesting activity, and studies by Kruglanski, Friedman, and Zeevi (1971) and Lepper, Greene, and Nisbett (1973) replicated the general finding, showing that other material and symbolic rewards could also undermine the intrinsic motivation of high school and preschool students.

Since those early studies, there has been continuing debate about the reliability and robustness of the findings and about what the findings mean in terms of the practical use of rewards as a motivational or behavior-change strategy. In this article, we review, evaluate, and interpret studies of the effects of rewards on intrinsic motivation.

Subsequent to the initial studies showing the undermining of intrinsic motivation by extrinsic rewards, a variety of additional studies were conducted that explored the effects of various types of rewards and other external events on intrinsic motivation. Several dozen studies focused on the effects of tangible rewards, many others explored the effects of positive feedback or *verbal rewards*, and still others moved beyond rewards to explore the effects of external events such as threats (Deci & Cascio, 1972), deadlines (Amabile, DeJong, & Lepper, 1976), competition (Deci, Betley, Kahle, Abrams, & Porac, 1981), and choice (Zuckerman, Porac, Lathin, Smith, & Deci, 1978). Additional studies also explored the effects of interpersonal contexts such as classroom climates (Deci, Schwartz, Sheinman, & Ryan, 1981), as well as the combined effects of events such as rewards and the interpersonal contexts within which they are administered (Ryan, Mims, & Koestner, 1983).

Although the ever-expanding field of research that began with exploration of the effects of extrinsic rewards on intrinsic motivation has moved in these numerous directions, the original finding of the undermining of intrinsic motivation by tangible extrinsic rewards has continued to be the focus of considerable controversy.

---

Edward L. Deci and Richard M. Ryan, Department of Clinical and Social Sciences in Psychology, University of Rochester; Richard Koestner, Department of Psychology, McGill University, Montreal, Quebec, Canada.

This work was supported by National Institute of Mental Health Grant MH-53385 and by a grant from the Social Science and Humanities Research Council of Canada.

We thank Liza Rovniak, Mireille Joussemette, and Laird Rawsthorne for their assistance with this project.

Correspondence concerning this article should be addressed to Edward L. Deci, Department of Clinical and Social Sciences in Psychology, University of Rochester, Rochester, New York 14627. Electronic mail may be sent to deci@psych.rochester.edu.

Articles spanning a period of 25 years have presented behaviorist perspectives maintaining that the "obscure" concept of intrinsic motivation inhibits "scientific progress" and that there is "no acceptable evidence" of an undermining effect (see, e.g., Carton, 1996; Flora, 1990; Scott, 1975). Nonetheless, many behavioral theorists (e.g., Bernstein, 1990), as well as motivational and cognitive theorists, have acknowledged that there is indeed a phenomenon of undermining to be reckoned with and have tried to specify the limiting conditions of the phenomenon and to provide theoretical accounts of the processes through which rewards can affect intrinsic motivation. Although the explanations are numerous, they tend to fall into three general categories: motivational theories, attributional theories, and behavioral or cognitive-behavioral theories. We consider these three types of theories in turn.

### Motivational Approaches

In line with the work of deCharms (1968) and Heider (1958), Deci (1971) suggested that an understanding of the effects of rewards requires a consideration of the interpretation (or what we call *functional significance*) that the recipients are likely to give to the rewards. Specifically, the recipients' interpretations of the rewards in relation to their own feelings of self-determination (i.e., autonomy) and competence were theorized to be the most important considerations. On the basis of the hypotheses and data from the Deci (1971) study and several that followed, Deci and Ryan (1980, 1985) made a formal statement of cognitive evaluation theory (CET).

CET asserts that underlying intrinsic motivation are the psychological needs for autonomy and competence, so the effects of an event such as a reward depend on how it affects perceived self-determination and perceived competence. Events that allow need satisfaction tend to increase intrinsic motivation whereas those that thwart need satisfaction tend to decrease intrinsic motivation. CET proposes that rewards can be interpreted by recipients primarily as controllers of their behavior or, alternatively, as indicators of their competence. In the former case, rewards are predicted to thwart satisfaction of the need for autonomy, lead to a more external perceived locus of causality (deCharms, 1968), and undermine intrinsic motivation. In the latter case, however, where rewards are positively *informational*, they are predicted to provide satisfaction of the need for competence and thus to enhance intrinsic motivation.

In many cases, rewards have conflicting effects, being experienced to some extent as controlling (thus thwarting satisfaction of the need for autonomy) and to some extent as informational (thus providing satisfaction of the need for competence). In such cases, these two processes work against each other, so additional factors must be taken into account in predicting the likely effect of such rewards. In considering such factors, we begin with an analysis of tangible rewards.

### Reward Contingencies

In making predictions about the effects of tangible rewards on intrinsic motivation for interesting activities, CET takes account of whether the rewards are expected while a person is doing a task and, if so, on what specific behaviors the rewards are made contingent. If rewards are not expected while the person is working

on a task, they are not predicted to affect intrinsic motivation for the task because the person is not likely to experience the task behavior as being controlled by the rewards.

For rewards that are expected while a person is working on a task, however, Ryan et al. (1983) introduced the following typology of reward contingencies: *task-noncontingent rewards*, which are given for something other than engaging in the target activity, such as simply participating in the study; *task-contingent rewards*, which are given for doing or completing the target activity; and *performance-contingent rewards*, which are given specifically for performing the activity well, matching some standard of excellence, or surpassing some specified criterion (e.g., doing better than 80% of the other participants). A further distinction can be usefully made between task-contingent rewards that are explicitly dependent upon completing the target task (herein referred to as *completion-contingent rewards*) and those that are dependent upon engaging in the activity but do not require completing it (herein referred to as *engagement-contingent rewards*). We have traditionally considered the completion-contingent and engagement-contingent rewards to constitute the single category of task-contingent rewards because the effects of these two reward contingencies have seemed to be remarkably similar; however, we separate them for this meta-analysis to evaluate quantitatively whether the effects of these two reward contingencies are in fact the same.

CET uses these reward contingencies as a factor in its analysis of the extent to which rewards tend to be interpreted as controllers of behavior versus affirmations of competence, and thus the extent to which they undermine versus enhance intrinsic motivation for interesting activities. First, consider task-noncontingent rewards. These rewards do not require doing the task, completing the task, or doing the task well, so there is no reason to expect these rewards to be experienced as either informational or controlling with respect to the task. Accordingly, intrinsic motivation for the task is predicted not to be affected by these rewards.

With engagement-contingent rewards, people have to work on the task to get the reward, so the reward is likely to be experienced as controlling the task behavior. Because the reward carries little or no competence affirmation, it is unlikely to increase perceived competence, so there is nothing to counteract the negative effects of the control. Thus, engagement-contingent rewards are predicted to undermine intrinsic motivation. With completion-contingent rewards, people have to complete the task to get the rewards, so the rewards are likely to be experienced as even more controlling. However, with completion-contingent rewards, receipt of the rewards implies some level of competence if the task required skill and the individual had a normative sense of what constitutes good performance on the task. To the extent that the rewards do represent competence affirmation, that implicit positive feedback could offset some of the control. Nonetheless, on average across different types of tasks, the competence-affirming aspect of completion-contingent rewards is not expected to be strong relative to the controlling aspect, so completion-contingent rewards are expected to undermine intrinsic motivation at a level roughly comparable to engagement-contingent rewards.

Finally, with performance-contingent rewards, where rewards are linked to people's performance, there is even stronger control—people have to meet some standard in order to maximize rewards—so there is a strong tendency for these rewards to un-

determine intrinsic motivation. However, performance-contingent rewards can also convey substantial positive competence information in cases where the person does well enough to get a level of reward that signifies excellent performance. In those cases, there is a significant tendency for performance-contingent rewards to affirm competence and, accordingly, to offset some of the negative effects of control.

Because of the competing tendencies of the controlling and competence-affirmation aspects of performance-contingent rewards, CET suggests that various other factors that affect the salience of these two aspects of the rewards need to be taken into account in making predictions about the rewards' effects on intrinsic motivation. Such factors include whether or not the level of reward implies excellent performance and whether or not the interpersonal climate within which the performance-contingent rewards are administered is demanding and controlling. We consider the issue of interpersonal context here and return later in the article to the issue of whether the rewards convey excellence.

### *Interpersonal Contexts*

In general, we use the term interpersonal context to refer to the social ambience of settings such as homes, classrooms, or work groups as they influence people's experience of autonomy, competence, or relatedness (Deci & Ryan, 1991; Ryan, 1995). With respect to the interpersonal context within which rewards are administered, the most important issue is the extent to which the ambience is controlling versus noncontrolling, in other words, the extent to which people within the context feel pressured to think, feel, or behave in particular ways. To a large extent, the controlling versus noncontrolling quality of an interpersonal context is determined by the style of administration used by the person administering the reward—for example, the teacher in a classroom or the manager in a work group (see, e.g., Deci, Connell, & Ryan, 1989; Deci, Nezlek, & Sheinman, 1981). When studied in laboratory experiments, the interpersonal climate is usually instantiated as the interpersonal style used by the experimenter to administer rewards, feedback, or other external events (see, e.g., Ryan, 1982; Ryan et al., 1983).

CET predicts that when the interpersonal style of administering performance-contingent rewards is relatively pressuring, the rewards tend to be experienced as more controlling, thereby leading to more diminishment of intrinsic motivation, whereas when the interpersonal style is relatively noncontrolling, the rewards tend to be experienced as more informational, thereby leading to less diminishment or possible enhancement of intrinsic motivation. For example, Ryan et al. (1983) compared two performance-contingent rewards groups, one in which the rewards were administered in a relatively controlling manner and the other in which they were administered in a relatively noncontrolling manner. As predicted, the controlling administration led to undermining of intrinsic motivation relative to the noncontrolling administration.

### *Symbolic Cue Value*

Harackiewicz and colleagues (Harackiewicz, 1979; Harackiewicz, Manderlink, & Sansone, 1984) introduced the concept of *cue value*, offering it as an additional element that could influence the effects of performance-contingent rewards on intrinsic moti-

vation. Cue value sensitizes individuals to the competence information in a situation and thus highlights the competence affirmation contained within performance-contingent rewards that are given for having done well. In essence, cue value is said to emphasize the informational aspect of performance-contingent rewards and to offset their controlling aspect. Accordingly, performance-contingent rewards were predicted and found to enhance intrinsic motivation relative to a condition in which participants were told that they would be evaluated and were subsequently given positive feedback but did not get a reward, because the reward (which is said to have the cue value) was not present to offset the control contained within the evaluation (Harackiewicz et al., 1984). However, performance-contingent rewards did not enhance intrinsic motivation relative to a no-reward control group in which participants received the positive feedback but had not been told they would be evaluated. In fact, in a study by Harackiewicz (1979), performance-contingent rewards that were not accompanied by norms led to significant undermining of intrinsic motivation relative to a control group that got the same positive feedback without rewards. Harackiewicz et al. (1984) interpreted this as indicating that normative information about people's performance needs to be available for performance-contingent rewards to have high cue value.

### *Verbal Rewards*

The term verbal rewards is not typically used in the current intrinsic motivation literature; instead, the term positive feedback is more commonly seen. We use the term verbal rewards here to make it easier to encompass the positive feedback studies within the general category of reward effects and thus to compare their effects with those of tangible rewards. In this regard, the original Deci (1971) article also included a study in which positive feedback was found to enhance intrinsic motivation.

According to CET, the informational aspect of verbal rewards is generally expected to be salient, and thus verbal rewards are generally predicted to enhance intrinsic motivation. However, verbal rewards can also have a significant controlling component—that is, people sometimes engage in behaviors to gain acknowledgment or approval—so verbal rewards also have the potential to undermine intrinsic motivation. CET suggests that, just as with tangible rewards, the interpersonal context within which positive feedback is administered influences how it is interpreted and thus what effect it has. Two studies have tested the prediction that controlling positive feedback leads to less intrinsic motivation than does informational positive feedback; both found strong support for the prediction. In one (Ryan, 1982), an example of controlling feedback was a statement containing the word *should*—namely, “Excellent, you should keep up the good work”—that was intended to be experienced as pressure to continue performing well. This was contrasted with feedback that simply provided the participants with their scores and with information indicating that they had done considerably above average. In the other study (Pittman, Davey, Alafat, Wetherill, & Kramer, 1980), an example of controlling feedback was “I haven't been able to use most of the data I have gotten so far, but you are doing really well, and if you keep it up I'll be able to use yours,” and an example of informational feedback was “Compared to most of my subjects, you are doing really well.” The idea in the controlling feedback was that the

experimenter's needing the participant to do well would be experienced as pressure by the participant. In both studies, the controlling feedback led to less intrinsic motivation than did the informational feedback.

### Attributional Approaches

Several investigators (e.g., Higgins & Trope, 1990; Kruglanski, 1975; Lepper, 1981) have provided attributional theories to interpret findings from the studies of rewards and intrinsic motivation. The best known of these (Lepper et al., 1973) built on Bem's (1972) self-perception theory to suggest an *overjustification effect*. People are said to make postbehavioral attributions about the causes of their own behavior based on a consideration of the behavior and the conditions within which it occurred. When people are rewarded for doing an interesting activity, they are likely to attribute their behavior to the reward and thus discount their interest in the activity as the cause of their behavior, leading to postbehavior intrinsic motivation that is lower than it would be if they had not gotten the reward.

Lepper, Sagotsky, Dafoe, and Greene (1982) expanded the attributional analysis because research indicated that young children do not operate with the discounting principle in making attributions but do show evidence of intrinsic motivation being undermined by rewards. These authors suggested that young children develop concrete social schemas about people's use of extrinsic incentives and sanctions to coerce behavior and that they gradually develop reflective generalizations such as the discounting principle from concrete social schemas. Lepper et al. provided an example of how, over time, children might develop the generalization "When people give me rewards so that I will do a task, the task is probably boring" from a set of more concrete experiences such as "When dad tells me I have to do something before I can go out to play, it is probably something I don't want to do."

Lepper (1981) proposed that the more salient the instrumentality between an activity and reward, the more likely the reward is to undermine intrinsic motivation, but when the reward conveys positive competence information, the information tends to offset some of the effect of the instrumentality so that the reward has a less detrimental effect on intrinsic motivation.

There is considerable similarity in the primary predictions of CET and the overjustification formulation, even though the theoretical processes are said to be different. First, both predict that salient rewards made contingent on doing a behavior undermine intrinsic motivation for that behavior. However, the overjustification approach focuses on the instrumentality per se, whereas CET focuses on the experience of feeling controlled, suggesting that not only the instrumentality but also other factors such as the interpersonal context can affect people's experience of control and thus their intrinsic motivation in situations where they are rewarded. Second, both approaches predict that when rewards convey positive competence information, they are less likely to undermine intrinsic motivation, although CET provides a more differentiated account of when and how the competence component is likely to operate.

### Behavioral Approaches

Some behaviorists such as Scott (1975) have asserted that the concept of intrinsic motivation is obscure and therefore not useful

and that there is no evidence supporting the undermining of intrinsic motivation. Others have claimed that the undermining effect (typically referred to as variability in response rates that might occur following reward termination) can be explained by standard operant principles (see, e.g., Carton, 1996; Flora, 1990). Still others (e.g., Bernstein, 1990; Dickinson, 1989; Eisenberger & Cameron, 1996) have not stuck strictly to operant principles but have attempted to explain the undermining effect "in terms comfortable to behavior analysts" (Bernstein, 1990, p. 331). In most of those discussions, however, although their authors have provided post hoc explanations for why responding following reward removal might have been below baseline, they have not attempted to apply behavioral principles in a comprehensive way to make differentiated predictions about the effects of different types of rewards and reward contingencies.

Dickinson (1989) offered three behavioral explanations for the undermining effect, including the following. Behaviors called intrinsic are, she said, culturally valued and reinforced by generalized praise. The fact that these behaviors are not linked to explicit reinforcers is part of the basis for approval. The presence of tangible rewards may therefore lead to less generalized praise, so when a tangible reward has been operative and is then removed, people are left with less generalized approval, which results in less free-choice behavior.

Bernstein (1990) expanded Skinner's (1945) analysis of how people learn to infer private events within a language community. Children use publicly available stimuli, and gradually, their verbal behavior describing private events is influenced by the same discriminative stimuli used by others. When their behavior is under external control (i.e., extrinsic rewards), self-descriptions of private events do not endorse the actions (hence, low intrinsic interest is reported).

Eisenberger and Cameron (1996) built on Seligman's (1975) learned helplessness theory to interpret intrinsic motivation results. They suggested that when people receive rewards that are not dependent on the quality of their performance (e.g., engagement-contingent and task-noncontingent rewards), they learn that they cannot control the receipt of rewards, so they become helpless. This helplessness, the authors suggested, has been misinterpreted by many researchers as decrements in intrinsic motivation. Further, they suggested, when rewards are dependent on the quality of performance (i.e., performance-contingent, completion-contingent, and verbal rewards), the rewards facilitate learned industriousness, which means that effortful activity takes on secondary reward properties, an effect that many have interpreted as increased intrinsic motivation.

The Eisenberger and Cameron argument is the most differentiated of the behavioral accounts in that it explicitly predicts that engagement-contingent and task-noncontingent rewards undermine intrinsic motivation, whereas performance-contingent, completion-contingent, and verbal rewards for effortful activity leave unchanged or enhance intrinsic motivation.

### Meta-Analytic Studies

The phenomenon of rewards undermining intrinsic motivation, as well as the CET and overjustification interpretations, has generated a great deal of interest, research, and debate. It is hardly surprising, therefore, that four meta-analyses have attempted to

evaluate the validity of various hypotheses within this field of research, particularly the hypothesis that expected tangible rewards, under most circumstances, undermine intrinsic motivation. Before reviewing those meta-analyses, we comment briefly on the meta-analytic procedure.

The use of meta-analyses to answer questions in areas with a large number of studies has substantial appeal because it provides a means of quantifying effect sizes and combining them across studies. This is particularly useful when researchers are attempting to isolate phenomena, as in inductive research, or when researchers are attempting to evaluate the effects of some fixed variable, such as the gender of participants. The method is less well suited to research areas that are driven by differentiated theories, and the research area of reward effects on intrinsic motivation is just such an area. Consider for a moment the evolution of the field. The first few studies, published in top-rated journals, showed that expected tangible rewards undermine intrinsic motivation. The politics of publishing in a theory-driven field demand that after a phenomenon has been replicated a few times, an additional study is unlikely to be published if it does not identify limiting conditions or extend theoretical understanding by specifying a moderator variable leading to an interaction. The meta-analytic method typically collapses across such interactions or unusual manipulations, particularly if there are only a few relevant studies. Thus, the method is inherently conservative in such theoretically guided research fields because important findings are obscured when they are collapsed.

That is precisely the case in this field. Numerous studies introduced moderators, finding decreased intrinsic motivation at some levels of the moderator variables and increased intrinsic motivation at other levels. For example, although the field has always been defined as an examination of the effects of rewards on intrinsic motivation for interesting activities, several studies introduced boring activities in the expectation of finding that the undermining effect did not hold for those activities (see, e.g., Daniel & Esser, 1980; Loveland & Olley, 1979; McLoyd, 1979; Newman & Layton, 1984). Examples of other moderators found to yield interactions, with rewards leading to an undermining of intrinsic motivation at one level of the moderator but not at another, included whether the rewards were informationally or controllingly administered (see, e.g., Ryan et al., 1983), whether the rewards were endogenous or exogenous to the target activity (see, e.g., Kruglanski et al., 1975), whether success at the activity was said to depend on ability versus luck (see, e.g., Weiner, 1980), whether there were norms about payment (see, e.g., Staw, Calder, Hess, & Sanderlands, 1980) or norms about performance (see, e.g., Harackiewicz, 1979), whether the rewards were salient or nonsalient (see, e.g., Ross, 1975), and whether positive versus negative feedback accompanied the rewards (see, e.g., Rosenfield, Folger, & Adelman, 1980; Salancik, 1975). In cases with a sufficient number of studies of a moderator variable, a meta-analysis can consider that variable, but in cases with an insufficient number, the data are typically collapsed across that variable, resulting in the conclusion of no effect for rewards in those studies when in fact there were highly significant effects that had been theoretically predicted.

In the four previous meta-analyses, as we show below, little attention was given to the moderator variables identified above. In our meta-analyses, we considered these moderators if there were enough studies, but when there were not enough studies, we too

collapsed across the important moderators and then discussed them in the text. Clearly, our collapsing across moderators means that the reported effects are weaker than the actual effects, particularly with respect to the undermining of intrinsic motivation by various types of tangible rewards, because it is the undermining effect that was explicitly intended to be delimited by studies of moderator variables. We turn now to a review of the previous meta-analyses.

### *Rummel and Feinberg*

The first, conducted by Rummel and Feinberg (1988), tested the CET hypothesis that extrinsic rewards with a salient controlling aspect undermine intrinsic motivation. Rummel and Feinberg reviewed 45 studies published between 1971 and 1985 and reported 88 effect sizes. Inclusion was explicitly limited to studies involving rewards that seemed to be controlling, so the authors did not evaluate more general questions such as whether all rewards, all tangible rewards, or all expected rewards undermine intrinsic motivation or whether different reward-contingencies have differential effects on intrinsic motivation. Their meta-analysis included studies of verbal as well as tangible rewards and expected as well as unexpected rewards without doing moderator analyses; it also included studies of deadlines (e.g., Amabile et al., 1976) and imposed goals (Mossholder, 1980), neither of which involved rewards.

Of the 88 effect sizes reported, 83 showed the undermining effect, and only 5 showed enhancement. Rummel and Feinberg (1988) concluded that "This meta-analysis lends support to the adequacy of [CET]" (p. 160).

Thus, the first meta-analysis supported both CET and the validity of an undermining effect. However, it did not do moderator analyses to examine issues such as the effects of different reward contingencies, it did not include unpublished doctoral dissertations, and it did not do separate analyses for the two dependent measures commonly used in this type of research, namely, target behavior during a free-choice period (see, e.g., Deci, 1972b) and self-reports of interest/enjoyment (Kruglanski, Alon, & Lewis, 1972).

### *Wiersma*

A subsequent meta-analysis by Wiersma (1992) included 20 studies published between 1971 and 1990 that used behavioral dependent measures. Of these, 11 used only the standard measure of free-time behavior, 4 used only task performance during the experimental phase, and 5 used both. In all, 28 effect sizes were reported. Although it has seldom been suggested that performance while the reward contingency is still in effect represents a measure of intrinsic motivation, Wiersma treated it as such and claimed to be evaluating whether extrinsic rewards undermine intrinsic motivation using these two measures.

Results from the 16 studies with a free-choice measure showed that rewards undermined intrinsic motivation, with an average effect size (Cohen's  $d$ ) of  $-0.50$ . Results also showed that rewards enhanced performance, with an average effect size of  $0.34$ , although that is not directly relevant to the issue being considered here, and other reviews have found results inconsistent with these (see, e.g., McGraw, 1978).

To summarize, although the results of Wiersma's meta-analysis showed an undermining of intrinsic motivation by extrinsic rewards and thus complemented those of Rummel and Feinberg, there were a number of shortcomings. With only 16 published studies of intrinsic motivation included, the meta-analysis omitted many relevant studies (including unpublished dissertations); some studies did not have no-reward control groups, instead comparing task-contingent rewards with task-noncontingent rewards; and reward contingencies were not examined separately.

### *Tang and Hall*

An additional meta-analysis was conducted by Tang and Hall (1995) to explore the overjustification effect. Tang and Hall reviewed 50 published studies involving 256 effect sizes that used six different dependent measures, including free-choice behavior and self-reported interest/enjoyment, as well as quantity of performance, quality of performance, ratings by others, and time waited to initiate the target activity during the free-choice period. Further, Tang and Hall evaluated specific hypotheses, including the hypotheses that expected task-contingent rewards would undermine intrinsic motivation and that positive feedback would enhance it, but they did not do comprehensive analyses including aggregate categories.

Their results showed that task-contingent rewards decreased intrinsic motivation ( $d = -0.51$ ), that performance-contingent rewards also undermined intrinsic motivation, ( $d = -0.35$ ), but that unexpected rewards did not affect intrinsic motivation ( $d = 0.12$ ). Results concerning verbal rewards and feedback were less clear, although there was indication that positive feedback did enhance intrinsic motivation ( $d = 0.34$ ). Finally, analyses of reward effects on uninteresting tasks suggested that there may be a positive effect on intrinsic motivation for uninteresting activities.

The Tang and Hall meta-analysis was more extensive and more sophisticated than the first two, and its results provided strong support for the proposition that expected tangible rewards undermine intrinsic motivation. However, it did not include unpublished dissertations, and the meaning of its results is somewhat obscured by the use of six different dependent measures, at least three of which were not pure measures of intrinsic motivation because they were taken during the reward phase rather than after rewards were terminated. These measures included the quality of performance during the reward phase (Brockner & Vasta, 1981), the quantity of performance during the reward phase (Eisenstein, 1985), and ratings of how much participants enjoyed the activity during the reward phase (Morgan, 1981). These measures reflected a mix of extrinsic and intrinsic motivation for the rewards group but only intrinsic motivation for the no-reward group, making them inappropriate for comparing the intrinsic motivation of the two groups.

### *Summary*

Three previous meta-analyses that focused on somewhat different issues all concluded that there is substantial support for the general hypothesis that expected tangible rewards made contingent upon doing, completing, or excelling at an interesting activity undermine intrinsic motivation for that activity. Although each of these meta-analyses had methodological shortcomings, the consistency in their results is noteworthy.

### *Eisenberger, Cameron, and Pierce*

Cameron and Pierce (1994) reported a hierarchical meta-analysis of reward effects that analyzed separately for free-choice behavior and self-reported attitudes. Later, this meta-analysis was slightly redone, using the same studies and the same effect sizes but different groupings of studies, and was published by Eisenberger and Cameron (1996) in *American Psychologist* with very little methodological detail provided. In referring to their meta-analyses in this article, we sometimes refer to one of the articles and sometimes to the other because, although essentially the same meta-analysis, the different articles reported different pieces of the methods and results.

Cameron and Pierce (1994) differentiated rewards as verbal versus tangible, tangible rewards as unexpected versus expected, and expected rewards as contingent on task completion or performance quality versus not contingent on completion or quality. Finally, the three contingency categories specified in the Ryan et al. (1983) typology (viz., task-noncontingent, task-contingent, and performance-contingent) were examined separately.

Cameron and Pierce (1994) reported no overall reward effect on free-choice behavior but significant enhancement on self-reported attitudes ( $d = 0.14$ ), significant enhancement by verbal rewards on both measures ( $d = 0.38$  and  $0.39$ , respectively), significant undermining by tangible rewards on free-choice behavior ( $d = -0.21$ ) but not on attitudes, no effect for unexpected tangible rewards on either measure, significant undermining by expected tangible rewards on free-choice behavior ( $d = -0.25$ ) but not on attitudes, no effect for contingent rewards (requiring task completion or high quality performance) on either measure, significant undermining by noncontingent rewards on free-choice behavior ( $d = -0.26$ ) but not on attitudes, no effect for task-noncontingent rewards on either measure, significant undermining by task-contingent rewards on free-choice behavior ( $d = -0.23$ ) but not on attitudes, and no effect for performance-contingent rewards on free-choice behavior but enhancement on attitudes ( $d = 0.19$ ). On the basis of these results, Cameron and Pierce concluded that there was no reason not to use reward systems, particularly as a motivational strategy in educational settings, and called for "abandoning cognitive evaluation theory" (p. 396).

Eisenberger and Cameron (1996) used reward-contingency categories that were different from Cameron and Pierce's (1994). They were *quality-dependent rewards*, which are given for meeting a performance standard; *completion-dependent rewards*, which are given for completing a task; and *performance-independent rewards*, which do not require completing the task or doing it well. In this new set of categories, the term quality-dependent rewards was simply a renaming of performance-contingent rewards. The completion-dependent rewards category resulted from Cameron and Pierce's having divided their task-contingent category into two subcategories—task-contingent rewards that are contingent according to a behavioral definition and task-contingent rewards that are noncontingent according to a behavioral definition. Completion-dependent rewards in the Eisenberger and Cameron typology were the ones referred to as task-contingent rewards that are contingent according to a behavioral definition in the Cameron and Pierce analysis. (They are comparable to what we call completion-contingent rewards in this article.) Eisenberger and Cameron did not make clear what exactly went into their perfor-

mance-independent category, which appears to have been simply a renaming of what Cameron and Pierce referred to as non-contingent rewards according to a behavioral definition, which were those that resulted from combining task-contingent rewards that are noncontingent according to a behavioral definition and task-noncontingent rewards. (In this article, we refer to those two groups as engagement-contingent and task-noncontingent, respectively.) However, Cameron and Pierce reported using 34 studies (with outliers removed) in their noncontingent category, whereas Eisenberger and Cameron reported using only 31 studies (with outliers removed) in their performance-independent category. It is thus possible that Eisenberger and Cameron simply excluded all task-noncontingent studies from their meta-analysis (there were only four task-noncontingent studies in the Cameron and Pierce list after they removed outliers). If Eisenberger and Cameron excluded all task-noncontingent studies, then their performance-independent category would be equivalent to our engagement-contingent category.

Eisenberger and Cameron (1996) reported that, after removing outliers, there was no overall reward effect on free-choice behavior but significant enhancement on attitudes ( $d = 0.14$ ), significant enhancement for verbal rewards on both measures ( $d = 0.38$  and  $0.39$ , respectively), significant undermining for tangible rewards on free-choice behavior ( $d = -0.21$ ) but not on attitudes, no effect for unexpected tangible rewards on either measure, significant undermining by expected tangible rewards on behavior ( $d = -0.25$ ) but not attitudes, no effect for quality-dependent rewards on behavior but significant enhancement on attitudes ( $d = 0.19$ ), no effect for completion-dependent rewards on either measure, and significant undermining for performance-independent rewards on behavior ( $d = -0.29$ ) but not attitudes. On the basis of these results, the authors claimed that the undermining effect was largely a myth.

Subsequent to the publication of Cameron and Pierce's (1994) meta-analysis, several articles (e.g., Kohn, 1996; Lepper, Keavney, & Drake, 1996; Ryan & Deci, 1996) criticized it for methodological inadequacies. The most important of these included (a) collapsing across cells where there were significant interactions for theoretically meaningful variables—for example, initial task interest, positive versus negative feedback, and informational versus controlling administration of rewards—without doing moderator analyses for any of those variables, (b) using inappropriate control groups in several comparisons, and (c) omitting nearly 20% of the studies as outliers rather than attempting to isolate the cause of the variability in effect sizes.

Clearly, although the results of the four meta-analyses were obtained by different methods and therefore cannot be compared directly, the Eisenberger and Cameron/Cameron and Pierce analysis (hereinafter Eisenberger, Cameron, and Pierce, unless reference to one of the original studies is being made) arrived at conclusions quite discrepant from those of the other three. It is therefore worth considering carefully the procedures employed by this group.

First, Eisenberger, Cameron, and Pierce included conditions from some studies using tasks the experimenters had selected specifically because they were dull and boring, and they collapsed the effects of rewards on interesting tasks and dull/boring tasks (see, e.g., Calder & Staw, 1975; Daniel & Esser, 1980; Hamner & Foster, 1975; McLoyd, 1979). As discussed earlier, this is problematic because the field of inquiry has always been defined in

terms of intrinsic motivation for interesting tasks, and the undermining phenomenon has always been specified as applying only to interesting tasks insofar as with boring tasks there is little or no intrinsic motivation to undermine. In the original research, the boring tasks were included primarily to examine limiting conditions to the undermining effect, so if one were to include the dull tasks, the principles of meta-analysis require that one analyze for moderating effects rather than simply collapsing across them (see, e.g., Johnson & Eagly, in press). Because early studies of this issue showed a crossover interaction (see, e.g., Calder & Staw, 1975; McLoyd, 1979), the type of task should have been analyzed as a moderator variable early in the hierarchical analysis in order not to obscure this important moderation at each level of analysis. As noted above, Eisenberger, Cameron, and Pierce also collapsed across several other moderator variables, thus obscuring significant effects and theoretically interesting issues.

Second, in the Eisenberger, Cameron, and Pierce meta-analyses, there were several instances in which inappropriate control groups were used. For example, Eisenberger, Cameron, and Pierce included a study by Boal and Cummings (1981) in which all groups of participants received monetary rewards; so there was not a no-reward control group.<sup>1</sup> As another example, in a study by

<sup>1</sup> In the Boal and Cummings (1981) study, individuals were hired through a temporary employment agency to work for 4 days, 4 hr per day, at \$3.00 per hour, at an obviously boring activity of calculating and transcribing data from a property rental record onto a coding form. At the beginning of the third day, some participants got an unexpected \$.25 raise ostensibly because their performance had been so good during the first 2 days. A second group got a cost-of-living raise, and a third group got no raise. In what they considered the critical comparison, the investigators compared the group that got the raise for having done well with the group that got no raise. They reported that this raise for good performance did not undermine intrinsic motivation. Apart from the fact that the task was in all likelihood experienced as uninteresting, the more important issues are, first, that there was not a group getting no monetary rewards to serve as a proper control group, and second, even if the raise group were considered a rewards group and the no-raise group a no-reward control group, the reward would have to be classified as an unexpected tangible reward relative to the first 2 days, which is the time that the reward group's performance was said to have led to their getting the raise. To examine the effect of expected rewards signifying good performance, participants must work on a task trying to do well enough to get a reward, but in the Boal and Cummings study, participants never tried to get the raise, so the study did not address the effect of pursuing rewards that imply good performance. As already noted, during the first 2 days, the subsequently rewarded group's members had no idea that they would get a raise if they did well (so, with respect to the first 2 days, the reward was unexpected), and furthermore, during the second 2 days, they were being paid on a hourly basis (\$3.25 per hour) and did not have to do well to keep getting their raised hourly rate. Thus, with respect to the second 2 days, the "rewards" were task-noncontingent. In fact, CET predicts, and the current meta-analysis confirms, that both unexpected tangible rewards and task-noncontingent tangible rewards do not affect intrinsic motivation, which is exactly what the results of the Boal and Cummings study show if one considers the comparison between the raise group and the no-raise group an appropriate test of the effects of rewards. Oddly, Cameron and Pierce (1994) classified the raise as a task-contingent reward that was contingent using the behavioral definition, which is essentially equivalent to what we call a completion-contingent reward, although there is no sense in which the raise could be construed as completion contingent.

Swann and Pittman (1977, Exp. 2) where one group of children got tangible rewards plus positive feedback but there was not a no-reward control group that got positive feedback, Eisenberger, Cameron, and Pierce included the reward plus feedback condition in their analysis, comparing it with the inappropriate no-reward, no-feedback control group. Additional studies for which they used inappropriate comparisons are pointed out in Appendix A, where we compare how we treated each study in our meta-analyses with how it was treated by Eisenberger, Cameron, and Pierce.

A third problem is that the Eisenberger, Cameron, and Pierce meta-analyses misclassified several studies. A study by Brennan and Glover (1980), for example, was listed by Cameron and Pierce (1994) as task-noncontingent, yet the method section made clear that the rewards were engagement-contingent—participants were given the rewards if they “work[ed] with the Soma puzzle for at least 8 minutes” (Brennan & Glover, 1980, p. 29). Differences in the way we versus Eisenberger, Cameron, and Pierce classified studies also appear in Appendix A.

A fourth problem is that Eisenberger, Cameron, and Pierce did not include in their list of studies some experiments that were published during the period covered by their meta-analyses and met their inclusion criteria, nor did they include unpublished doctoral dissertations. Also, as specified in Appendix A, some of the experiments included in their list of studies were not used in some of the analyses where they were appropriate.

### *Overall Summary of Previous Meta-Analyses*

Variability among the methods and approaches of the four meta-analytic studies makes it difficult to compare and contrast the findings, although there appears to be consistency among the results of the Rummel and Feinberg (1988); Wiersma (1992), and Tang and Hall (1995) analyses, with only the Eisenberger and Cameron (1996) and Cameron and Pierce (1994) meta-analyses reporting different conclusions. In general, the most consistent finding was that task-contingent tangible rewards (i.e., those given for either engaging in or completing a task) were detrimental to intrinsic motivation. Two of the meta-analyses also found that verbal rewards enhanced intrinsic motivation, and there was the suggestion that undermining by tangible rewards was considerably stronger for free-choice behavior than for self-reports.

Differences in the results of the previous meta-analyses, significant methodological problems that were identified in each, failure of each to include unpublished doctoral dissertations, considerable controversy regarding appropriate conclusions from each, and the importance of this issue for real-world applications all argue that a more definitive meta-analysis evaluating the effects of extrinsic rewards on intrinsic motivation is warranted.

### **Method**

We used a hierarchical approach to conduct two primary meta-analyses, one for well-controlled experiments that used free-choice persistence as the dependent measure of intrinsic motivation and the other for well-controlled experiments that used self-reported interest. In each case, we began by examining the effects of all rewards on intrinsic motivation for interesting tasks and then differentiated reward conditions as follows. If a set of effects was heterogeneous, we divided the category using concepts that have been identified theoretically or empirically as possible moderators of reward effects. Only after we had exhausted the potential moderators did we

exclude any effects as outliers to produce homogeneity. Fewer than 4% of the effects were outliers, and we report analyses both with and without those outliers.

In carrying out this hierarchical procedure, we were guided primarily by the theoretical and empirical issues that have emerged as important considerations in the literature of this field. Thus, we began by examining verbal rewards (i.e., positive feedback) versus tangible rewards. Tangible-reward studies were examined in terms of unexpected versus expected rewards, and studies of expected tangible rewards were separated according to whether the rewards were task-noncontingent, engagement-contingent, completion-contingent, or performance-contingent. In addition, for each of the differentiated categories based on reward type and contingency type (viz., verbal, unexpected, task-noncontingent, engagement-contingent, completion-contingent, and performance-contingent), when the set of effects was heterogeneous, we did moderator analyses for age effects, comparing studies with children versus college students. Because this moderator variable showed significant effects in some differentiated categories, we also analyzed for whether it was a moderator variable at the aggregate levels of all verbal rewards and all tangible rewards. Finally, in the free-choice behavior analysis of the performance-contingent rewards category, to achieve homogeneity, we analyzed for differences among four groups defined by whether the control group got feedback or no feedback and by the valence of the feedback implicit in or accompanying the rewards.

For purposes of comparison with other meta-analyses, we also calculated composite effect sizes for task-contingent rewards. Finally, we did four supplemental analyses that seemed important. The first compared the effects of both verbal rewards and tangible rewards in published studies versus unpublished doctoral dissertations. The second, done to examine whether reward effects are only transitory, compared studies in which free-choice behavior was assessed immediately after the reward was terminated, versus within a week, versus more than a week later. The third compared reward effects for interesting versus dull tasks when both were included in the same experiment, and we then repeated the primary meta-analyses with the dull-task conditions included to ascertain whether the significant effects remained. The fourth supplemental analysis tested the CET prediction of how informational versus controlling administration of rewards moderated their effects on intrinsic motivation.

### *Identifying Relevant Studies*

We began by searching the PsycINFO and ERIC databases and recent editions of the Current Contents database for studies published between 1971 and August 1997. Varying combinations of the following search terms were entered: intrinsic motivation, rewards, reinforcement, free time, and free choice. Whenever possible, each of these search terms was entered as both keyword and subject heading to maximize the number of articles retrieved. We then reviewed the lists of studies used in each previous meta-analysis, adding studies that we had not previously identified, and reviewed the reference list of each article, thereby finding a small number of additional articles that had not been identified in the previous steps.

In searching the PsycINFO database, we looked for unpublished doctoral dissertations as well as for published studies. We identified 41 dissertations. Abstracts were reviewed to determine whether the dissertations included an experimental manipulation of rewards and an assessment of intrinsic motivation. Twenty-five dissertations appeared to satisfy these requirements and were ordered from Dissertation Abstracts International, although 1 of them ultimately had to be obtained from a university in Canada. Of the 25, 1 was not available, the results from 1 had appeared in a later published article, and 4 others failed to meet inclusion criteria. Thus, 19 dissertations were used, and 1 of these (Liberty, 1986) included two experiments, making a total of 20 dissertation studies that explored the effects of rewards. Of these, 18 used a free-choice measure, and 15 used a self-report measure. We applied the following inclusion requirements to all

identified studies from both published articles and unpublished dissertations.

### *Inclusion Criteria*

The studies examined in this meta-analysis were all experiments with at least one experimental and one control group, and all were done either in the laboratory or under well-controlled, laboratory-like conditions. A very small number of field studies that had been identified were excluded because of being noncomparable with the laboratory experiments.

### *Initial Task Interest*

Because the field of research being evaluated concerns reward effects on people's intrinsic motivation for interesting activities, we included in the primary meta-analyses only studies, or conditions within studies, in which the interest value of the target tasks was at least neutral. We used two criteria. First, if there was a measure of task interest reported for the control group (or for a pretreatment assessment of the rewards group) and the average on that measure was below the midpoint of the scale for the target task, it was considered an uninteresting task and not included in the meta-analysis. Two studies were excluded on this basis (Freedman & Phillips, 1985; Phillips & Freedman, 1985).<sup>2</sup> Alternatively, if the experimenter specifically defined the task as uninteresting, as was the case with Overskeid and Svartdal (1998), it was excluded. Further, in several studies, initial task interest was experimentally manipulated, with one task selected by the experimenter as interesting and one as uninteresting. We excluded the uninteresting tasks from the primary meta-analyses. Thirteen studies from 10 published articles and 3 unpublished dissertations manipulated task interest, including an interesting task and a boring task (Calder & Staw, 1975; Chung, 1995; Daniel & Esser, 1980; Eisenstein, 1985; Griffith, 1984; Hamner & Foster, 1975; Hitt, Marriot, & Esser, 1992; Loveland & Olley, 1979; McLoyd, 1979; Mynatt et al., 1978; Newman & Layton, 1984; Smith, 1980; Wilson, 1978). In a supplemental meta-analysis of these 13 studies, we examined the hypothesis of an interaction for initial task interest, with tangible rewards undermining intrinsic motivation for interesting tasks but not for uninteresting tasks.

If a study included only one condition that was considered an interesting task and did not have an average interest rating below the midpoint of an interest scale, it was included. If a study included two or more target tasks but these were not explicitly defined as interesting versus uninteresting, we included both, collapsing across the task conditions. For example, Arkes (1979) used easy and difficult tasks, and we included both in the analyses of interesting tasks; Tripathi and Agarwal (1988) used a heuristic and an algorithmic task, and we included both.

From the time of the first study that included interesting and dull tasks (Calder & Staw, 1975) as a way to examine limiting conditions to the undermining effect, there was indication that rewards affected the two kinds of task differently. Thus, including both types of task in a meta-analysis of a field that has been defined as the effects of extrinsic rewards on intrinsic motivation for interesting tasks would create problems such as obscuring the important findings of the field. Even though one could separate the effects of rewards on interesting versus uninteresting tasks in a moderator analysis, important results would still be obscured if separation of effects by that moderator variable was not done as the first step in the analysis. Furthermore, CET was explicitly formulated to explain the effects of external events such as rewards on intrinsic motivation for interesting tasks. If an activity stimulates little or no initial intrinsic motivation, one would not expect to find external events undermining intrinsic motivation, and the principles of CET would not apply. Instead, with boring tasks, the critical theoretical issue is how to facilitate internalization of the regulation of such tasks if they are deemed important within the social environment (see, e.g., Deci, Eghrari, Patrick, & Leone, 1994; Ryan & Connell, 1989). From our perspective, the theoretical principles relevant to motivation for

dull-boring tasks are different from those relevant to motivation for interesting tasks.

### *Dependent Measures*

The primary meta-analyses included only studies in which the assessment of intrinsic motivation was done subsequent to the experimental period when the reward contingency was no longer in effect (in a few cases where there was also a pretreatment baseline, we calculated effect sizes based on change scores if sufficient information was available to do so). Thus, we excluded measures of performance or attitudes taken while the reward contingency was still in effect because those measures reflected a mix of intrinsic and extrinsic motivation. As noted earlier, Tang and Hall (1995) had included such measures in their analysis.

The primary measure of intrinsic motivation, introduced by Deci (1971), is the degree to which participants return to and persist at the target activity during a free-choice period subsequent to the experimental phase. The assessment of free-choice persistence in more than 90% of the studies concerned the time spent on the target activity during the free-choice period. In all but 3 of those studies, the data were presented as the average number of seconds, but in the remaining 3, free-choice play was presented as the proportion of participants who spent any time with the target activity (Pallak, Costomiris, Sroka, & Pittman, 1982; Swann & Pittman, 1977, Exps. 1 and 2). Finally, in 6 published studies and 3 dissertation studies (Harackiewicz et al., 1984, Exps. 2 and 3; Lee, 1982; Liberty, 1986, Exps. 1 and 2; Pittman, Cooper, & Smith, 1977; T. W. Smith & Pittman, 1978; Weiner, 1980; Weiner & Mander, 1978), the amount of time was not directly assessed, but instead, free-choice persistence was assessed as the number of trials or successes with the target activity during the free-choice period. If a study reported both amount of time and performance during the free-choice period, we used only amount of time. One study that assessed both the amount of free-choice time and the number of trials during the free-choice period reported a correlation between the two of  $r = .94$  (Harackiewicz et al., 1984, Exp. 1). In the Cameron and Pierce (1994) meta-analysis, 5 of the 6 published studies that used only the performance measure (viz., Harackiewicz et al., 1984, Exps. 2 and 3; T. W. Smith & Pittman, 1978; Weiner, 1980; Weiner & Mander, 1978) were included in a separate meta-analysis of free-choice performance; Cameron and Pierce reported that the effects of rewards on free-choice time were virtually identical to the effects of rewards on free-choice performance. Thus, to be

<sup>2</sup> Within the industrial-organizational psychology literature, there have been a few experiments published that explored the effects of rewards on intrinsic motivation for uninteresting tasks in a worklike setting. Some of them manipulated initial task interest by adding some element to the task to make it more interesting. Doing that got the interest level of the "interesting" task above the midpoint for some studies (e.g., Hamner & Foster, 1975), and in those cases, we included the interesting task conditions in our primary meta-analysis and used both the interesting and uninteresting conditions in the supplemental meta-analysis that explored the moderation of reward effects by initial task interest (see *Supplemental Analyses*, subsection Interesting versus Boring Tasks, below). In the Phillips and Freedman (1985) study, the task was proofreading, and initial interest was varied by having participants read material from a law review in the boring condition and from short stories in the interesting condition. The mean for task interest (on a Likert 1-7 scale) was 2.55 for the boring task and 2.87 for the interesting task, indicating that the so-called interesting task was indeed nearly as boring as the task labeled boring and well below the midpoint of the scale. The Freedman and Phillips (1985) study reported very similar means for the interest value of the two proofreading tasks, on the same scale. Thus, these 2 studies do not meet our inclusion criterion of having tasks with interest levels of at least the midpoint on an interest scale.

inclusive, we used the 9 studies that reported only free-choice performance in our examination of reward effects on free-choice intrinsic motivation.<sup>3</sup> All 9 of these studies examined tangible rewards, and none examined verbal rewards, so to be sure that the use of these studies did not affect our results, we analyzed for whether tangible rewards had a different effect on free-choice time and free-choice performance. We found that they did not,  $Q_b(1) = 2.62, ns$ .

The second measure of intrinsic motivation was self-reported interest in the target activity, which was assessed either with a single item or with multiple items that were sometimes factor-analytically derived. In some cases, questions assessing interest and enjoyment were mixed together as a single factor, so we used the total of those items; in other cases, only enjoyment or liking of the task was assessed, so we used that. Several studies included assessments of other affective states such as satisfaction, but we did not include them unless they were part of a composite and interest was not reported separately. In one case (Harackiewicz, 1979), the only dependent variable reported was a composite of three self-report variables and one behavioral variable, so we included that composite in the self-report analyses but not the free-choice analyses; in another (S. Anderson & Rodin, 1989), the only dependent variable reported was a composite of the self-report and behavioral variables, and we again included that composite only in the self-report analyses. In part, we did so because Eisenberger, Cameron, and Pierce had used the composite measure from both of those studies in only their self-report analysis, so our also doing that seemed reasonable for the purpose of comparing the results of our meta-analysis with theirs. Further, Harackiewicz (personal communication, January 6, 1999) reported that, although the data are no longer available for analysis, the results had been comparable on each of the dependent measures, so the results on the composite measure were a good indicator of the results for self-reported interest.

### No-Reward Control Groups

We included only studies that had appropriate no-reward control groups. Thus, we excluded a few studies that Wiersma (1992) had included that compared task-contingent rewards to task-noncontingent rewards (see, e.g., Farr, 1976), and we excluded one study that Eisenberger, Cameron, and Pierce had included that compared one monetary-reward group that got a small unexpected raise with another monetary-reward group that did not get the raise (Boal & Cummings, 1981). For more detail on the Boal and Cummings study, see Footnote 1.

### Total Number of Studies Included

Using these criteria, the primary meta-analyses included 128 studies taken from 94 published articles and 19 dissertations. One hundred and one studies included a free-choice behavior measure of intrinsic motivation, and 84 included a self-report measure of interest/enjoyment. One additional study (Ryan, 1982) was also used in the supplemental meta-analysis of informational versus controlling reward administrations.

### Coding Studies

Each rewards condition was first coded for whether the rewards were tangible or verbal. Symbolic rewards such as "good-player certificates" were considered tangible even though positive feedback was typically implicit in them. Tangible rewards were then coded for whether they were expected or unexpected. In most cases of unexpected tangible rewards, the variable was experimentally manipulated to contrast unexpected with expected tangible rewards (see, e.g., Lepper et al., 1973). The great majority of tangible-reward studies involved expected rewards, which were further coded as task-noncontingent, engagement-contingent, completion-contingent, or performance-contingent, in accord with what they were contingent upon.

Task-noncontingent rewards are those given without specifically requiring the person to engage in the target activity. For example, Ross, Karniol, and Rothstein (1976) had a task-noncontingent reward condition in which participants were told "that if they waited for the experimenter to return, they would receive the candies as a reward" (p. 444). In that case, the reward was contingent on waiting, but it was not contingent on task engagement, so it was considered task-noncontingent. As another example, Deci (1972a) gave rewarded participants \$2 for being in the study, which was also considered task-noncontingent because the rewards were given for being in the study, not for doing the task. Of course, one can never tell whether participants assume that task-noncontingent rewards require doing the task, but these rewards should be less likely to be experienced by participants as the reason they are doing the task than would be the case with the other three reward contingencies.

Engagement-contingent rewards are specifically contingent upon engaging in the target task but do not require completing the task, doing it well, or reaching some standard. The Ross et al. (1976) study also included an engagement-contingent reward condition in which the children were told they would receive candies "if they would draw some pictures" (p. 444). Drawing, of course, was the target activity in that study. Another example is the so-called task-contingent conditions in Harackiewicz (1979) in which college students worked on hidden-figure puzzles for a specified amount of time. They did not have to complete the puzzles to get the rewards; indeed, there was no way for them to tell if they had finished a puzzle.

Completion-contingent rewards are made explicitly dependent upon completing a task. For example, Deci (1971) offered participants \$1 for each of four puzzles they completed within a specified time. The puzzles were relatively difficult, so completing the puzzles provided implicit positive feedback, although neither normative information nor explicit feedback was provided by the experimenter in either the tangible-rewards group or no-rewards group. Other studies required completing tasks that did not require as much skill, so the reward carried less implicit positive feedback. For example, Calder and Staw (1975) gave participants 15 puzzles to solve and told them that they would get a \$1 reward "when you finish" (p. 601). Because those puzzles did not require much skill, successful completion of the 15 puzzles did not provide much competence affirmation. As a final example, Fabes (1987, Exps. 1 and 2) had preschool children build something of their own choosing with blocks, and they were told that they could have a reward "when they finished building with the blocks" (p. 8). The children had to complete the construction to receive the reward, but the completion provided little competence affirmation because in those conditions there was not a right way to do it and thus no standard against which the children could implicitly compare their performance.

Performance-contingent rewards are offered dependent upon the participants' level of performance. This was the most complex category in terms of the different methods employed. One complexity was that, in some studies, performance-contingent rewards groups were compared with no-feedback control groups, whereas in other studies they were compared with control groups that got the same feedback that was implicit in the performance-contingent reward. For example, in a study by Boggiano and Ruble (1979), children in the performance-contingent reward group were told that they would get a reward if they obtained a score of at least 3, so getting the reward implied that they had done well, but nothing was said to the children in the control group about standards and no feedback was provided either explicitly or implicitly. In contrast, Ryan et al. (1983) did a

<sup>3</sup> The difference between our performance measure of trials during the free-choice period and the performance measures used in the Tang and Hall (1995) meta-analysis that we criticized is that Tang and Hall included performance measures taken during the reward phase, whereas those that we used were all taken during the free-choice period after the rewards had been terminated. As already noted, performance during the reward phase mixes extrinsic and intrinsic motivation for the rewarded participants.

study in which performance-contingent rewards were given to participants for doing well, and they were compared with no-reward control-group participants who were given the same feedback—that is, who were told that they had done well.

Furthermore, the type of implicit feedback conveyed by the performance-contingent rewards differed across studies. First, consider studies that used no-feedback control groups. In many cases, a specific performance-contingent reward was given to every participant for performing well or surpassing some stated criterion (e.g., doing better than 80% of other participants), thus signifying that each participant had done well. For example, Orlick and Mosher (1978) told children that "if you do a good job today and tomorrow on the balance board you will get this good balancer trophy" (p. 31), and each child in the reward condition did receive the trophy. Some performance-contingent rewards were, however, given for levels of performance that were not necessarily very good and may have varied across participants. For example, in a study by Pittman et al. (1977), college students could win between 5¢ and 25¢ on each of 10 trials of a game, depending on how well they did on that trial. Some participants received a relatively small amount, signifying poorer performance, whereas others received a larger amount, signifying better performance. Consider another example of performance-contingent rewards that conveyed a message of less than maximal performance. Daniel and Esser (1980) did a performance-contingent study with a no-feedback control group in which the rewarded participants were told that they could earn up to \$2 but were then given \$1, implying that their performance was less than optimal, perhaps at about the 50th percentile. There is a considerable difference in the meaning of performance-contingent rewards to participants if they receive the maximum amount (which signifies excellent performance) versus some amount less than maximum (which signifies less than maximum performance).

Now consider performance-contingent studies where the same feedback (whether implicit or explicit) was received by both the rewards group and the no-rewards group. Some studies varied the feedback, such that some participants (both rewarded and nonrewarded) got positive feedback indicating that they had done very well, and others got negative feedback indicating that they had done poorly. Thus, there were instances in which some people were told that they had done well and were compared with people who received a reward signifying that they had done well, and there were other instances in which some people were told that they had done badly and were compared with people who got a reward signifying that they had done badly. The meaning of a reward that conveys positive feedback is likely to be quite different from the meaning of a reward that conveys negative feedback. For example, in a study by Rosenfield et al. (1980), one group of participants got rewards for being in the top 15% of participants, and another group simply got feedback conveying that they were in the top 15%. Thus, the feedback was very positive for both groups. However, another group of participants in this study got rewards for being in the bottom 15%, and still another group simply got feedback conveying that they were in the bottom 15%. In actuality, the amount of the rewards was the same whether participants were in the positive feedback group or the negative feedback group.

In a second round of coding, all studies that had been coded as performance-contingent were further coded into one of four categories based on the above considerations. The categories were (a) no-feedback control group, maximum reward; (b) no-feedback control group, less than maximum reward; (c) feedback control group with positive feedback conveyed to both groups; and (d) feedback control group with negative feedback conveyed to both groups.

Although it may seem that placement of studies (or conditions within studies) into contingency categories would be relatively straightforward, there were instances where it was not. The particular contingency being studied (e.g., performance-contingent) was explicitly named and examined in some studies, but many studies were done to examine issues other than contingency categories, and in some of those, there was not an exact

specification of what participants were told they had to do to get the reward. Further, many studies were reported as studies of "task-contingent" rewards, but as already noted, that category included both engagement-contingent and completion-contingent rewards, and the study methods often did not specify exactly what participants were told about what the reward was for. Thus, judgment was often required. Consider for example the two studies in the Fabbes (1987) article. In the methods for Experiment 1, the specific instructions given to the children were presented, and it was explicit that the children would get rewards for finishing the construction. Thus, the rewards were completion-contingent. In the methods of Experiment 2, the exact wording of what the children were told was not provided, but the article said that children "were offered a reward for their participation" and that this "was the procedure used in Experiment 1" (p. 12). The phrase "for their participation" suggests that the rewards in Experiment 2 were engagement-contingent, but the statement that the method was the same as in Experiment 1, where the exact wording to the children was provided, indicates that the rewards were in fact completion-contingent. Thus, we coded Experiment 2 as completion-contingent also.

Because of the potential difficulty in coding studies into the contingency categories, coding was performed independently by the second author and a trained research assistant. Interrater reliability was assessed using Cohen's kappa ( $K$ ) which reflects agreement between judges, corrected for chance (J. Cohen, 1960). All reliabilities were well above  $K = .70$ , which is the minimum acceptable reliability (Bakeman & Gottman, 1989):  $K = .95$  for tangible versus verbal,  $K = .83$  for expected versus unexpected,  $K = .82$  for coding of the specific form of contingency (i.e., task-noncontingent, engagement, completion, performance), and  $K = 1.00$  for classification of the performance-contingent reward studies (or conditions within studies) into the four categories. Disagreements were resolved through discussion.

### Effect Sizes

In these meta-analyses, Cohen's  $d$  was used as the measure of effect size. It reflects the difference between the means of two groups divided by the pooled within-group standard deviations, adjusted for sample size (Hedges & Olkin, 1985). Separate effect sizes were calculated for behavioral and self-report measures of intrinsic motivation and were entered into separate meta-analyses. The mean of the control group was subtracted from the mean of the rewards group, so a negative  $d$  reflects an undermining effect in which the rewards group showed less intrinsic motivation than did the control group, whereas a positive  $d$  reflects an enhancement effect in which the rewards group showed more intrinsic motivation than the control group.

Means, standard deviations,  $t$  tests,  $F$  tests, and sample sizes were used to calculate  $d$  values, which by definition are corrected for sample size. When standard deviations were not reported, we estimated them from means and significance tests. For one study (Loveland & Olley, 1979),  $d$  was calculated directly from a significance level. For any study where insufficient data were provided to calculate an effect size, we assigned an effect of  $d = 0.00$  and included those imputed values in all analyses reported in this article. We assigned  $d = 0.00$  for 8 out of 150 free-choice effects and 26 out of 121 self-report effects, and we noted these effects in the appropriate tables. All effect-size computations and summary analyses were done using DSTAT (Johnson 1993), a meta-analytic software program.

Composite effect-size estimates ( $d+$ ) were the average of individual effects ( $d$ ) weighted by the reciprocal of their variance, thus giving greater weight to more reliable effect-size estimates (Hedges & Olkin, 1985). In this article, we typically use the notation  $d$  rather than  $d+$  when referring to composite  $d$  values, yet in all cases the composite  $d$  values are corrected for reliability. Each calculation of  $d+$  provides both a test of whether the value differs significantly from 0.00 and a 95% confidence interval (CI). Because the significance tests and CIs are largely redundant, we generally present only CIs. Only in cases where a CI has 0.00 at one end do we also

present the significance level, because the 0.00 has been rounded to two decimal places. The homogeneity of each set of effect sizes was tested by the within-class goodness-of-fit statistic ( $Q_w$ ), which has an approximate chi-square distribution with  $k - 1$  degrees of freedom, where  $k$  equals the number of effect sizes (Johnson, 1993). A significant  $Q_w$  statistic indicates systematic variation among effect sizes and suggests moderator variables (Johnson & Eagly, in press).

Each effect size we report within a particular category includes all relevant conditions. For example, a study by Orlick and Mosher (1978) included three reward groups (verbal, unexpected, and performance-contingent) and a no-reward control group. The individual effect size for each type of reward was entered in the relevant analysis (verbal, unexpected, and performance-contingent). The performance-contingent reward was also entered into the expected reward analysis. Further, an effect size was calculated for the two tangible-rewards conditions (using the one control group compared with the two relevant rewards groups) and entered in the analysis of tangible rewards, and an effect size across all three reward conditions was similarly calculated and entered in the global analysis of all types of rewards.

The statistical differences among different categories of reward effects were tested by computing the between-class goodness-of-fit statistic ( $Q_b$ ), which has an approximate chi-square distribution with  $p - 1$  degrees of freedom (where  $p$  equals the number of classes). A significant  $Q_b$  statistic indicates that the magnitude of the effect differs between classes of the moderator variable.

## Results

We begin by examining the aggregate effect of all reward types on intrinsic motivation. In this and all other analyses, the results for the free-choice behavioral measure are presented first, followed by the results for self-reported interest. We then proceed to differentiate categories of reward effects based on both theoretical and empirical considerations in an attempt to isolate moderators and to identify meaningful reward-effect categories that are homogeneous.

### *The Primary Meta-Analyses*

In presenting the two primary meta-analyses, the six tables of effect sizes for particular studies are presented at the level of the differentiated categories of reward and contingency types: verbal rewards, unexpected tangible rewards, task-noncontingent rewards, engagement-contingent rewards, completion-contingent rewards, and performance-contingent rewards. In each case, the table lists free-choice and self-report data in separate columns, and there is a column with a code for whether the participants were children (coded 1) or college students (coded 2). Most of the children were preschool or elementary students, but 6 studies used high school students (Harackiewicz, 1979; Harackiewicz, Abrahams, & Wagaman, 1987; Harackiewicz & Manderlink, 1984; Kruglanski et al., 1971, 1975, Exps. 1 and 2). In each table, studies that are unpublished dissertations have the indicator D after the name and date of the study. Finally, in the table for performance-contingent rewards, a code is used to denote the four categories, based on type of control group and feedback. If a performance-contingent study had conditions that fit in two of these different categories, the effect size for each category is listed separately in that table.

### *All Rewards*

*Free-choice behavior.* One hundred one studies of reward effects included a free-choice behavioral measure, and their com-

posite effect size was  $d = -0.24$  (CI =  $-0.29$ ,  $-0.19$ ), thus indicating that with all studies considered, rewards did significantly undermine intrinsic motivation for interesting activities. The set of effect sizes was heterogeneous, however,  $Q_w$  (100) = 287.62,  $p < .0001$ .

*Self-reported interest.* Eighty-four studies used a self-report measure, and their composite effect size was  $d = 0.04$  (CI =  $-0.02$ ,  $0.09$ ), indicating that with all relevant studies considered, rewards did not significantly affect self-reported interest for interesting activities. As with the free-choice measure, this set of effect sizes was heterogeneous,  $Q_w$  (83) = 205.17,  $p < .0001$ .

*Relation to previous meta-analyses.* Only the Eisenberger, Cameron, and Pierce meta-analyses reported the average effect sizes for all types of verbal and tangible rewards considered together. With 61 studies for free-choice, with outliers and  $d$  values of zero included, their average effect size was a nonsignificant  $d = -0.06$ , and with 64 effect sizes for self-reports, their corresponding effect size was a significant enhancement of  $d = 0.14$ . Those meta-analyses, of course, had considerably fewer studies than did ours, and their inclusion of studies without appropriate control groups and of experimental conditions with tasks that were selected by the experimenter to be dull and boring also accounts for some of the differences in results. Finally, various errors in their methods, which are detailed in Appendix A, account for further differences.

### *Verbal Rewards Versus Tangible Rewards*

Given the heterogeneity in both the free-choice and self-report meta-analyses for all reward effects, we tested whether verbal versus tangible rewards were a significant moderator, as both theoretical and empirical considerations suggest that these reward categories should have different effects. For the free-choice measure, the effects of verbal rewards and tangible rewards were significantly different,  $Q_b$  (1) = 90.19,  $p < .0001$ ; for the self-report measure, the effects of verbal rewards and tangible rewards were also significantly different,  $Q_b$  (1) = 43.87,  $p < .0001$ . Thus, we consider the effects of verbal rewards first and then the effects of tangible rewards.

### *Verbal Rewards*

*Free-choice behavior.* Twenty-one studies, listed in Table 1, explored the effects of verbal rewards (i.e., positive feedback) on free-choice behavior. The composite  $d = 0.33$  (CI =  $0.18$ ,  $0.43$ ) indicates that positive feedback did enhance intrinsic motivation. However, the set of 21 effects was heterogeneous,  $Q_w$  (20) = 32.71,  $p < .05$ , so we tested for a difference between the effects of verbal rewards on the free-choice behavior of children versus college students. There was a significant difference between the two age groups,  $Q_b$  (1) = 5.14,  $p < .02$ . The 7 studies of verbal-reward effects for children yielded a nonsignificant composite effect size of  $d = 0.11$  (CI =  $-0.11$ ,  $0.34$ ), and these effects were homogeneous,  $Q_w$  (6) = 6.77, ns. The 14 college-student effects were also homogeneous,  $Q_w$  (13) = 20.79, ns, with  $d = 0.43$  (CI =  $0.27$ ,  $0.58$ ). This is a particularly noteworthy finding because it suggests that, although verbal rewards enhance intrinsic motivation for college students, they do not do so for children.

**Table 1**  
*Effects of Verbal Rewards on Free-Choice Behavior and Self-Reported Interest, Presented as Cohen's *d*, Corrected for Sample Size*

| Study                            | Free-choice<br><i>d</i> | <i>n<sub>e</sub></i> , <i>n<sub>c</sub></i> <sup>a</sup> | Sample <sup>b</sup> | Self-report<br><i>d</i> |
|----------------------------------|-------------------------|--|---------------------|-------------------------|
| R. Anderson et al. (1976)        | 0.65                    | 18, 46   | 1                   |                         |
| S. Anderson & Rodin (1989)       |                         | 10, 10   | 2                   | 0.40                    |
| Blanck et al. (1984), Exp. 1     | 0.55                    | 70, 69   | 2                   | 0.69                    |
| Blanck et al. (1984), Exp. 2     | 0.70                    | 12, 12   | 2                   | 0.00 <sup>c</sup>       |
| Butler (1987)                    |                         | 50, 50   | 1                   | 1.41                    |
| D. S. Cohen (1974) D             | 0.11                    | 13, 13   | 2                   | 0.11                    |
| Crino & White (1982)             | 0.01                    | 20, 10   | 2                   | 0.07                    |
| Danner & Lonky (1981), Exp. 2    | -0.08                   | 30, 30   | 1                   | -0.07                   |
| Deci (1971), Exp. 3              | 0.79                    | 12, 12   | 2                   | 0.00 <sup>c</sup>       |
| Deci (1972b)                     | 0.29                    | 48, 48   | 2                   |                         |
| Deci et al. (1975)               | 0.02                    | 32, 32   | 2                   |                         |
| Dollinger & Thelen (1978)        | -0.07                   | 12, 12   | 1                   | 0.00 <sup>c</sup>       |
| Efron (1976) D                   |                         | 13, 15   | 2                   | 0.98                    |
| L. W. Goldstein (1977) D         | 0.41                    | 16, 16   | 1                   | -0.08                   |
| Harackiewicz (1979)              |                         | 31, 31   | 1                   | 0.58                    |
| Kast & Connor (1988)             |                         | 90, 30   | 1                   | -0.46                   |
| Koestner et al. (1987)           | 0.50                    | 35, 18   | 2                   | 0.00 <sup>c</sup>       |
| Orlick & Mosher (1978)           | -0.34                   | 11, 12   | 1                   |                         |
| Pallak et al. (1982)             | -0.10                   | 30, 12   | 1                   |                         |
| Pittman et al. (1980)            | 0.78                    | 24, 12   | 2                   |                         |
| Pretty & Seligman (1984), Exp. 1 | 0.14                    | 10, 10   | 2                   | 0.67                    |
| Ryan et al. (1983)               | 0.52                    | 32, 16   | 2                   | 0.00 <sup>c</sup>       |
| Sansone (1986)                   |                         | 44, 11   | 2                   | 0.67                    |
| Sansone (1989)                   |                         | 82, 41   | 2                   | 0.46                    |
| Sansone et al. (1989)            |                         | 40, 40   | 2                   | 0.12                    |
| Shanab et al. (1981)             | 0.63                    | 20, 20   | 2                   | 0.42                    |
| W. E. Smith (1975) D             | 0.07                    | 24, 24   | 2                   | 0.00 <sup>c</sup>       |
| Tripathi & Agarwal (1985)        | 1.58                    | 20, 20   | 2                   | 0.47                    |
| Vallerand (1983)                 |                         | 40, 10   | 1                   | 1.95                    |
| Vallerand & Reid (1984)          |                         | 28, 28   | 2                   | 0.42                    |
| Zinser et al. (1982)             | 0.08                    | 64, 32   | 1                   |                         |

Note. D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup> The sample size for the experimental condition (*n<sub>e</sub>*) is reported first, followed by the sample size for the control group (*n<sub>c</sub>*). <sup>b</sup> In this column, 1 signifies that the participants were children, and 2 signifies that they were college students. <sup>c</sup> In these cases, *d* = 0.00 was assigned to the study because its authors failed to report relevant cell means.

**Self-reported interest.** Twenty-four verbal-reward studies, also listed in Table 1, included a self-report measure. The composite *d* = 0.36 (CI = 0.25, 0.48) indicates that positive feedback did enhance interest. However, these effects were heterogeneous, *Q<sub>w</sub>* (23) = 75.13, *p* < .0001. An analysis of age effects yielded a nonsignificant result, *Q<sub>b</sub>* (1) = 0.18, ns. Thus, 3 outliers were excluded from the group of all verbal-rewards self-report studies to achieve homogeneity (Butler, 1987; Kast & Connor, 1988; Vallerand, 1983). The results for the 21 homogeneous effects, *Q<sub>w</sub>* (20) = 22.34, ns, showed significant enhancement, *d* = 0.31 (CI = 0.19, 0.44).

Together, the meta-analyses of verbal rewards show that verbal rewards enhanced both free-choice behavior and self-reported interest, but further analyses indicate that the enhancement in free-

choice intrinsic motivation by verbal rewards applied only to college students and not to children.

**Relation to previous meta-analyses.** Tang and Hall (1995), Eisenberger and Cameron (1996), and Cameron and Pierce (1994) investigated the effects of positive feedback (verbal rewards) and, like us, all found significant enhancement of intrinsic motivation with average effect sizes in the small to moderate range (J. Cohen, 1988). However, none of the previous meta-analyses analyzed for age effects, so this important new finding was previously obscured. In sum, although the various meta-analyses used different studies, the results are generally consistent for the overall effects of verbal rewards.

### Tangible Rewards

**Free-choice behavior.** When all 92 tangible-rewards studies were considered together, the effect was highly significant, *d* = -0.34 (CI = -0.39, -0.28), showing undermining of the free-choice behavior by tangible rewards. However, this set of studies was heterogeneous, *Q<sub>w</sub>* (91) = 225.70, *p* < .0001. Because both theoretical and empirical considerations suggest that expected and unexpected tangible rewards may have different effects, we tested for a difference between these two categories and found a significant difference on free-choice intrinsic motivation, *Q<sub>b</sub>* (1) = 10.62, *p* < .001. Thus, we analyzed these two categories separately in subsequent analyses.

**Self-reported interest.** Seventy studies of tangible-rewards assessed self-reported interest. The composite effect size for tangible rewards showed significant undermining relative to no rewards, *d* = -0.07 (CI = -0.13, -0.01). This set of studies was also found to be heterogeneous, *Q<sub>w</sub>* (69) = 127.43, *p* < .0001. An analysis of the difference between expected and unexpected rewards effects did not yield a significant difference for self-reports, *Q<sub>b</sub>* (1) = 0.87, ns. However, for theoretical reasons and to maintain parallel analyses for free-choice and self-report data, we proceeded in the same way as for free-choice by analyzing the effects of unexpected and expected tangible rewards separately.

In sum, the analyses indicate that the effect of all tangible rewards was to significantly undermine intrinsic motivation, assessed with both measures.

**Relation to previous meta-analyses.** Cameron and Pierce (1994) calculated the effect size for all tangible rewards, and their results were also reported in Eisenberger and Cameron (1996). For 43 studies involving free-choice behavior, they concluded that there was significant undermining by all tangible rewards (*d* = -0.21), and for 33 studies with self-reports, they concluded that there was no effect for tangible rewards. Thus, their free-choice results, like ours, showed undermining for all tangible rewards, although they included fewer studies and their average effect size was considerably smaller than ours. For self-reports, we found significant undermining, but they reported no effect.

### Unexpected Tangible Rewards

One of the earliest findings in the intrinsic motivation literature was that rewards introduced after task performance, such that participants did not expect them while working on the target activity, tended not to affect intrinsic motivation for the activity (Lepper et al., 1973). Studies that examined this effect are summarized in Table 2.

**Table 2**  
*Effects of Unexpected Tangible Rewards on Free-Choice Behavior and Self-Reported Interest, Presented as Cohen's *d*, Corrected for Sample Size*

| Study                                 | Free-choice<br><i>d</i> | <i>n<sub>e</sub></i> , <i>n<sub>c</sub></i> <sup>a</sup> | Sample <sup>b</sup> | Self-report<br><i>d</i> |
|---------------------------------------|-------------------------|--|---------------------|-------------------------|
| Eisenstein (1985)                     | 0.33                    | 10, 10   | 1                   |                         |
| Greene & Lepper (1974)                | 0.12                    | 29, 15   | 1                   |                         |
| Harackiewicz et al. (1984),<br>Exp. 2 | 0.43                    | 15, 15<br>36, 33   | 2<br>1              | 0.15<br>-0.57           |
| Kruglanski et al. (1972)              |                         |  |                     |                         |
| Lepper et al. (1973)                  | 0.11                    | 18, 15   | 1                   |                         |
| Orlick & Mosher (1978)                | -1.30                   | 12, 12   | 1                   |                         |
| Pallak et al. (1982)                  | -0.36                   | 15, 12   | 1                   |                         |
| Petty & Seligman (1984),<br>Exp. 1    | 0.06                    | 30, 30   | 2                   | 0.40                    |
| Petty & Seligman (1984),<br>Exp. 2    | 0.06                    | 30, 30   | 2                   | 0.38                    |
| W. E. Smith (1975) D                  | 0.09                    | 24, 24   | 2                   | 0.00 <sup>c</sup>       |

*Note.* D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup>The sample size for the experimental condition (*n<sub>e</sub>*) is reported first, followed by the sample size for the control group (*n<sub>c</sub>*). <sup>b</sup>In this column, 1 signifies that participants were children, and 2 signifies that they were college students. <sup>c</sup>In this case, *d* = 0.00 was assigned to the study because its author failed to report relevant cell means.

**Free-choice behavior.** Nine studies involved unexpected tangible rewards, and they yielded an average *d* = 0.01 (CI = -0.20, 0.22), indicating, as predicted, that unexpected tangible rewards do not affect free-choice behavior. These studies were homogeneous, *Q<sub>w</sub>* (8) = 11.54, ns.

**Self-reported interest.** Five of the studies also included self-report measures, and they yielded a *d* = 0.05 (CI = -0.19, 0.29), thus also indicating no effect. This set of effects was also homogeneous, *Q<sub>w</sub>* (4) = 10.03, ns.

**Relation to previous meta-analyses.** Both Tang and Hall (1995) and Eisenberger and Cameron (1996) analyzed the effects of unexpected tangible rewards, and they too found that unexpected rewards do not affect intrinsic motivation.

### Expected Tangible Rewards

**Free-choice behavior.** The set of 92 expected-tangible-rewards studies yielded a composite *d* = -0.36 (CI = -0.42, -0.30), thus indicating that expected tangible rewards significantly undermined free-choice intrinsic motivation, with a moderate effect size. This set of expected-tangible-reward effects was, however, heterogeneous, *Q<sub>w</sub>* (91) = 224.85, *p* < .0001. An analysis of differences among the four contingency categories (task-noncontingent, engagement-contingent, completion-contingent, and performance-contingent) did not yield a significant effect, *Q<sub>b</sub>* (3) = 5.97, *p* < .12, but we nonetheless analyzed these four categories separately. Both theory and research suggest that the reason an expected tangible reward is given (i.e., upon what it is made contingent) may influence the reward's effect. Task-noncontingent rewards—that is, rewards not requiring task behavior—should be less likely to affect intrinsic motivation for the task than are rewards requiring engagement with, completion of, or high-quality performance on the task. In fact, an analysis did show

that task-noncontingent effects differed from engagement-contingent effects,  $\chi^2 = 3.67, p < .05$ .<sup>4</sup>

**Self-reported interest.** Sixty-nine studies included a self-report measure and yielded a significant composite *d* of -0.07 (CI = -0.13, -0.01). The set of effects was heterogeneous, *Q<sub>w</sub>* (68) = 121.60, *p* < .0001. An analysis for differences among the four contingency types was significant, *Q<sub>b</sub>* (3) = 8.84, *p* < .05, so these studies were further analyzed in terms of reward contingency. Further analyses indicated that task-noncontingent effects differed from engagement-contingent effects,  $\chi^2 = 5.56, p < .05$ , and that engagement-contingent effects differed from performance-contingent effects,  $\chi^2 = 4.58, p < .05$ .

In sum, the meta-analyses indicate that expected tangible rewards significantly undermine intrinsic motivation, assessed both behaviorally and through self-reports.

**Relation to previous meta-analyses.** The previous meta-analysis by Wiersma (1992) included only expected tangible rewards and reported an average *d* of -0.50, which is a bit stronger than our result. Eisenberger and Cameron (1996) also included a summary analysis of expected tangible rewards and found a significant undermining of *d* = -0.25 with 42 studies of free-choice behavior. With 32 studies involving self-report, they reported no significant effect for expected tangible rewards.

### Task-Noncontingent Rewards

Nine studies fell into the task-noncontingent category and are listed in Table 3. A typical example of task-noncontingent rewards is the Deci (1972a) study where participants were paid \$2 simply for participating in the experiment, independent of what they did while there. The idea was to simulate the typical work condition in which people are paid a salary without having their pay tied in any direct way to the specific tasks they perform.

**Free-choice behavior.** The 7 studies of task-noncontingent rewards with a free-choice measure yielded a nonsignificant composite *d* = -0.14 (CI = -0.39, 0.11). This set of effects was homogeneous, *Q<sub>w</sub>* (6) = 5.91, ns.

**Self-reported interest.** Five of the studies also had self-report data, and the composite effect size for them was also nonsignificant, *d* = 0.21 (CI = -0.08, 0.50), and homogeneous, *Q<sub>w</sub>* (4) = 9.19, ns.

**Relation to previous meta-analyses.** Two of the previous meta-analyses examined the effects of task-noncontingent rewards. With 31 effect sizes, Tang and Hall (1995) found a nonsignificant effect (*d* = 0.12) using their six different dependent measures. With four effect sizes, Cameron and Pierce (1994) found nonsignificant effects for the free-choice measure (*d* = 0.10) and the self-report measure (*d* = -0.01). Thus, the three meta-analyses support the conclusion that task-noncontingent rewards tend not to affect intrinsic motivation.

### Engagement-Contingent Rewards

This category includes rewards given for engaging in a task, independent of whether the task is completed or done well. A

<sup>4</sup>The chi-square statistic is used within DSTAT to do follow-up tests of specific comparisons between groups after the *Q<sub>b</sub>* omnibus test has been done for more than two groups.

**Table 3**  
*Effects of Task-Noncontingent Tangible Rewards on Free-Choice Behavior and Self-Reported Interest, Presented as Cohen's *d*, Corrected for Sample Size*

| Study                             | Free-choice<br><i>d</i> | <i>n<sub>e</sub></i> , <i>n<sub>c</sub></i> <sup>a</sup> | Sample <sup>b</sup> | Self-report<br><i>d</i> |
|-----------------------------------|-------------------------|--|---------------------|-------------------------|
| Dafoe (1985) D                    | -0.23                   | 24, 24   | 1                   | 0.68                    |
| Deci (1972a)                      | 0.01                    | 24, 16   | 2                   |                         |
| Earn (1982)                       | -0.28                   | 40, 20   | 2                   | 0.18                    |
| Kruglanski et al. (1971)          |                         | 16, 16   | 1                   | -0.69                   |
| Okano (1981), Exp. 2              | -0.84                   | 11, 11   | 1                   | 0.24                    |
| Pittman et al. (1982),<br>Exp. 1  | 0.23                    | 10, 10   | 1                   |                         |
| Ross et al. (1976)                | 0.48                    | 12, 12   | 1                   |                         |
| Swann & Pittman (1977),<br>Exp. 1 | -0.20                   | 20, 20   | 1                   |                         |
| Wimperis & Farr (1979)            |                         | 16, 16   | 2                   | 0.45                    |

Note. D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup>The sample size for the experimental condition (*n<sub>e</sub>*) is reported first, followed by the sample size for the control group (*n<sub>c</sub>*). <sup>b</sup>In this column, 1 signifies that the participants were children, and 2 signifies that they were college students.

typical example is the condition labelled task-contingent in the Ryan et al. (1983) study, in which participants were told that they would receive a reward for engaging in a series of hidden-figures puzzles. Because participants did not know how many figures were hidden in each drawing, they did not know whether they had completed or done well at the task. The engagement-contingent category included more studies than any other specific reward-contingency category, and all these studies are listed in Table 4.

*Free-choice behavior.* The 55 studies with a free-choice measure yielded a composite effect size of  $d = -0.40$  ( $CI = -0.48, -0.32$ ), indicating clearly that engagement-contingent rewards undermine intrinsic motivation. This set of effect sizes was heterogeneous,  $Q_w(54) = 143.89, p < .001$ , so we compared the effects of engagement-contingent rewards administered to children versus college students. The analysis revealed a significant difference between the two age groups,  $Q_b(1) = 6.76, p < .01$ . The set of 43 studies with children yielded a  $d = -0.46$  ( $CI = -0.55, -0.37$ ), but it was heterogeneous,  $Q_w(42) = 118.85, p < .001$ , so 4 outliers were removed (Boggiano, Ruble, & Pittman, 1982; Danner & Lonky, 1981, Exp. 2; Morgan, 1983, Exp. 1; Swann & Pittman, 1977, Exp. 2). The resulting set of 39 effect sizes yielded a composite  $d = -0.43$  ( $CI = -0.53, -0.34$ ) and was homogeneous,  $Q_w(38) = 51.09, ns$ . The 12 college-student studies, which yielded a  $d = -0.21$  ( $CI = -0.37, -0.05$ ), were homogeneous,  $Q_w(11) = 18.27, ns$ . To summarize, the results indicate that engagement-contingent rewards undermined free-choice intrinsic motivation for both children and college students, but the composite effect size for children was significantly greater than the composite effect size for college students.

It is interesting to note that 2 of the children studies came from the Ross (1975) article in which he investigated the effects of salience versus nonsalience of the rewards as a moderator. In the above analyses, we collapsed across those conditions; however, when we looked at them separately, we found that reward salience did moderate reward effects. Specifically, when the 2 studies of

free-choice behavior were combined meta-analytically, the composite effect size for salient rewards was  $d = -0.78$  ( $CI = -1.59, -0.30$ ), whereas the composite effect size for nonsalient rewards was  $d = 0.24$  ( $CI = -0.23, 0.71$ ). Thus, if only salient rewards had been used, as they were in nearly all other studies in the meta-analysis, the effects on free-choice intrinsic motivation of all engagement-contingent rewards, and especially the effects of engagement-contingent rewards with children, would have been even stronger.

*Self-reported interest.* The 35 engagement-contingent studies with a self-report measure yielded a composite  $d = -0.15$  ( $CI = -0.25, -0.06$ ), indicating that engagement-contingent rewards undermine self-reported intrinsic interest. This set of effects was homogeneous,  $Q_w(34) = 38.28, ns$ .

*Relation to previous meta-analyses.* The meta-analyses by Rummel and Feinberg (1988), Wiersma (1992), and Tang and Hall (1995) did not examine this category of rewards, although both Wiersma and Tang and Hall examined task-contingent rewards, which consist of the combination of engagement-contingent and completion-contingent rewards. That category will be examined later. Eisenberger and Cameron (1996) had a category called performance-independent rewards, which was either equivalent to our engagement-contingent category or was a combination of engagement-contingent and task-noncontingent, but, as noted earlier, they did not define it clearly enough for us to tell. They reported significant undermining by performance-independent rewards for free-choice ( $d = -0.29$ ) with 31 effects, but not for self-reports ( $d = 0.08$ ) with 14 effects.

#### Completion-Contingent Rewards

This category includes rewards that are given for completing one or more tasks. The first published intrinsic motivation experiment (Deci, 1971) was an example of a completion-contingent rewards study: Participants worked on four variations of a three-dimensional puzzle called SOMA and were given \$1 for each puzzle they completed in the allotted time.

*Free-choice behavior.* The 20 completion-contingent studies that used a free-choice measure are listed in Table 5. The composite effect size for these 20 studies was  $d = -0.36$  ( $CI = -0.50, -0.22$ ). However, the set of effects was heterogeneous,  $Q_w(19) = 35.02, p < .01$ . We then examined for age differences, but the comparison was not significant,  $Q_b(1) = 2.09, ns$ . We removed one outlier (Deci, 1972b), and the resulting  $d = -0.44$  ( $CI = -0.59, -0.30$ ) for 19 studies was homogeneous,  $Q_w(18) = 23.81, ns$ . The results therefore indicate that completion-contingent rewards significantly undermine free-choice behavior.

*Self-reported interest.* The 15 completion-contingent studies with a self-report measure are also listed in Table 5. The composite effect size was  $d = -0.03$  ( $CI = -0.18, 0.13$ ), but these effects were heterogeneous,  $Q_w(14) = 43.96, p < .001$ . An analysis for age effects was not significant,  $Q_b(1) = 2.00, ns$ , so 2 outliers (Kruglanski et al., 1975, Exp. 1; Wimperis & Farr, 1979) were removed to produce homogeneity,  $Q_w(12) = 16.81, ns$ . The average  $d = -0.17$  ( $CI = -0.33, -0.00$ ) for the 13 effects was significant ( $p < .05$ ). Completion-contingent rewards were thus found to undermine intrinsic motivation, assessed with both measures.

It is worth noting that the Kruglanski et al. (1975) study that had

Table 4  
*Effects of Engagement-Contingent Rewards on Free-Choice Behavior and Self-Reported Interest,  
 Presented as Cohen's *d*, Corrected for Sample Size*

| Study                            | Free-choice <i>d</i> | <i>n<sub>e</sub>, n<sub>c</sub></i> <sup>a</sup> | Sample <sup>b</sup> | Self-report <i>d</i> |
|----------------------------------|----------------------|--|---------------------|----------------------|
| Amabile et al. (1986), Exp. 1    | 0.00 <sup>c</sup>    | 56, 57   | 1                   | 0.00 <sup>c</sup>    |
| Amabile et al. (1986), Exp. 3    |                      | 30, 30   | 2                   | 0.00 <sup>c</sup>    |
| R. Anderson et al. (1976)        | -0.53                | 36, 19   | 1                   |                      |
| Arnold (1976)                    |                      | 17, 36   | 2                   | 0.00 <sup>c</sup>    |
| Arnold (1985)                    |                      | 13, 16   | 2                   | -0.03                |
| Boggiano & Ruble (1979)          | -0.94                | 20, 20   | 1                   |                      |
| Boggiano et al. (1982)           | 0.28                 | 81, 84   | 1                   |                      |
| Boggiano et al. (1985)           | -0.65                | 13, 13   | 1                   |                      |
| Brennan & Glover (1980)          | 0.52                 | 20, 19   | 2                   |                      |
| Brewer (1980) D                  | -0.20                | 24, 24   | 1                   | 0.12                 |
| Chung (1995)                     | -1.15                | 5, 5   | 1                   |                      |
| Danner & Lonky (1981), Exp. 2    | -1.27                | 30, 30   | 1                   | -1.15                |
| DeLoach et al. (1983)            | 0.00 <sup>c</sup>    | 26, 26   | 1                   |                      |
| Dimitroff (1984) D               | -0.28                | 108, 18  | 1                   | -0.26                |
| Efron (1976) D                   |                      | 12, 13   | 2                   | 0.17                 |
| Fabes et al. (1986)              | 0.05                 | 24, 24   | 1                   | -0.14                |
| Fabes et al. (1988)              | -1.25                | 14, 14   | 1                   | -0.69                |
| Fabes et al. (1989)              | -0.77                | 15, 14   | 1                   |                      |
| Greene & Lepper (1974)           | -0.69                | 15, 14   | 1                   |                      |
| Griffith (1984) D                | -0.73                | 42, 42   | 1                   |                      |
| Hamner & Foster (1975)           |                      | 15, 15   | 2                   | -0.15                |
| Harackiewicz (1979)              |                      | 31, 31   | 1                   | -0.56                |
| Hitt et al. (1992)               | -0.67                | 30, 15   | 2                   | -0.48                |
| Hyman (1985) D                   | -0.44                | 32, 32   | 1                   |                      |
| Karniol & Ross (1977)            | -0.08                | 17, 20   | 1                   |                      |
| Lepper et al. (1973)             | -0.65                | 18, 15   | 1                   |                      |
| Lepper et al. (1982), Exp. 3     | -0.50                | 32, 32   | 1                   |                      |
| Loveland & Olley (1979)          | -1.17                | 6, 6   | 1                   |                      |
| Morgan (1981), Exp. 1            | -0.97                | 27, 27   | 1                   | -0.14                |
| Morgan (1981), Exp. 2            | -0.72                | 20, 20   | 1                   | 0.04                 |
| Morgan (1983), Exp. 1            | -1.92                | 40, 40   | 1                   | -0.55                |
| Morgan (1983), Exp. 2            | -0.66                | 20, 20   | 1                   | 0.00 <sup>c</sup>    |
| Mynatt et al. (1978)             | -0.11                | 5, 5   | 1                   |                      |
| Newman & Layton (1984)           | -0.36                | 21, 10   | 1                   |                      |
| Ogilvie & Prior (1982)           | -0.08                | 26, 26   | 1                   |                      |
| Okano (1981), Exp. 1             | -0.96                | 15, 15   | 1                   | -0.44                |
| Okano (1981), Exp. 2             | -1.22                | 10, 11   | 1                   | 0.00                 |
| Patrick (1985) D                 | 0.00 <sup>c</sup>    | 33, 31   | 1                   | -0.23                |
| Perry et al. (1977)              | -0.43                | 32, 32   | 1                   | -0.17                |
| Picek (1976) D                   | 0.00 <sup>c</sup>    | 10, 10   | 2                   | -0.65                |
| Pittman et al. (1982), Exp. 1    | 0.12                 | 10, 10   | 1                   |                      |
| Pittman et al. (1982), Exp. 2    | -0.16                | 27, 27   | 1                   |                      |
| Pretty & Seligman (1984), Exp. 1 | -0.77                | 30, 30   | 2                   | -0.06                |
| Pretty & Seligman (1984), Exp. 2 | -0.11                | 30, 30   | 2                   | -0.16                |
| Reiss & Sushinsky (1975), Exp. 1 | -0.82                | 16, 16   | 1                   |                      |
| Ross (1975), Exp. 1              | -0.18                | 40, 20   | 1                   | -0.45                |
| Ross (1975), Exp. 2              | -0.65                | 52, 14   | 1                   | 0.00 <sup>c</sup>    |
| Ross et al. (1976)               | -0.68                | 12, 12   | 1                   |                      |
| Ryan et al. (1983)               | -0.35                | 16, 16   | 2                   | 0.00 <sup>c</sup>    |
| Sarafino (1984)                  | -0.41                | 85, 15   | 1                   | 0.00 <sup>c</sup>    |
| Shiffman-Kaufman (1990) D        | 0.06                 | 20, 20   | 1                   | -0.05                |
| A. T. Smith (1980) D             | -0.75                | 21, 27   | 1                   |                      |
| Swann & Pittman (1977), Exp. 1   | -0.79                | 20, 20   | 1                   |                      |
| Swann & Pittman (1977), Exp. 2   | -1.31                | 26, 26   | 1                   |                      |
| Thompson et al. (1993)           | 0.14                 | 34, 33   | 2                   |                      |
| Tripathi & Agarwal (1985)        | 0.29                 | 20, 20   | 2                   | 0.52                 |
| Tripathi & Agarwal (1988)        | 0.00 <sup>c</sup>    | 20, 10   | 2                   | 0.67                 |
| Weiner & Mander (1978)           | -0.31                | 30, 30   | 2                   | 0.00 <sup>c</sup>    |
| Williams (1980)                  | -0.31                | 24, 24   | 1                   | 0.00 <sup>c</sup>    |
| Wilson (1978) D                  | -0.34                | 23, 23   | 2                   | -0.06                |
| Yuen (1984) D                    | -0.40                | 60, 60   | 2                   | -0.12                |

Note. D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup>The sample size for the experimental condition (*n<sub>e</sub>*) is reported first, followed by the sample size for the control group (*n<sub>c</sub>*). <sup>b</sup>In this column, 1 signifies that the participants were children, and 2 signifies that they were college students. <sup>c</sup>In these cases, *d* = 0.00 was assigned to the study because its authors failed to report relevant cell means.

**Table 5**  
*Effects of Completion-Contingent Rewards on Free-Choice Behavior and Self-Reported Interest, Presented as Cohen's  $d$ , Corrected for Sample Size*

| Study                               | Free-choice<br>$d$ | $n_e, n_c^a$ | Sample <sup>b</sup> | Self-report<br>$d$ |
|-------------------------------------|--------------------|--------------|---------------------|--------------------|
| Arkes (1979)                        | -0.16              | 32, 32       | 2                   | 0.03               |
| Arnold (1985)                       |                    | 13, 16       | 2                   | -0.05              |
| Boggiano et al. (1985)              | -0.86              | 13, 13       | 1                   |                    |
| Brockner & Vasta (1981)             | -0.36              | 26, 26       | 2                   | -0.57              |
| Calder & Staw (1975)                |                    | 10, 10       | 2                   | -0.76              |
| Deci (1971), Exp. 1                 | -0.51              | 12, 12       | 2                   | 0.00 <sup>c</sup>  |
| Deci (1972b)                        | 0.33               | 64, 32       | 2                   |                    |
| Eisenstein (1985)                   | -0.50              | 18, 10       | 1                   |                    |
| Fabes (1987), Exp. 1                | -0.77              | 18, 19       | 1                   |                    |
| Fabes (1987), Exp. 2                | -0.41              | 14, 14       | 1                   |                    |
| Feehan & Enzle (1991),<br>Exp. 2    | -0.95              | 24, 12       | 2                   |                    |
| G. S. Goldstein (1980) D            | -0.30              | 14, 14       | 2                   |                    |
| L. W. Goldstein (1977) D            | -0.96              | 16, 16       | 1                   | -0.84              |
| Griffith et al. (1984)              | 0.00 <sup>c</sup>  | 64, 32       | 1                   |                    |
| Hamner & Foster (1975)              |                    | 20, 15       | 2                   | -0.26              |
| Kruglanski et al. (1975),<br>Exp. 1 |                    | 24, 24       | 1                   | 1.10               |
| Liberty (1986), Exp. 1 D            | -0.84              | 31, 24       | 2                   | -0.32              |
| Liberty (1986), Exp. 2 D            | -0.24              | 42, 42       | 2                   | 0.03               |
| Luyten & Lens (1981)                | -0.98              | 10, 10       | 2                   | -1.22              |
| McGraw & McCullers<br>(1979)        |                    | 18, 17       | 2                   | -0.04              |
| McLoyd (1979)                       | -1.01              | 18, 9        | 1                   |                    |
| Porac & Meindl (1982)               | -0.76              | 20, 20       | 2                   |                    |
| Sorensen & Maehr (1976)             | -0.65              | 20, 20       | 1                   |                    |
| Staw et al. (1980)                  |                    | 47, 46       | 2                   | 0.17               |
| Vasta & Stirpe (1979)               | -0.16              | 4, 5         | 1                   |                    |
| Weiner (1980)                       | 0.20               | 24, 24       | 2                   | 0.00 <sup>c</sup>  |
| Wimperis & Farr (1979)              |                    | 16, 16       | 2                   | 1.28               |

Note. D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup>The sample size for the experimental condition ( $n_e$ ) is reported first, followed by the sample size for the control group ( $n_c$ ). <sup>b</sup>In this column, 1 signifies that the participants were children, and 2 signifies that they were college students. <sup>c</sup>In these cases,  $d = 0.00$  was assigned to the study because its authors failed to report relevant cell means.

a  $d = 1.10$  and was removed as an outlier was a study of whether exogenous versus endogenous monetary rewards differentially affect intrinsic motivation. Exogenous rewards, which are not an inherent part of the task, are the type of rewards used in virtually all the studies in this meta-analysis, and they actually showed a substantial undermining of interest ( $d = -0.87$ ), whereas endogenous rewards, which are inherent in the task (in Kruglanski et al., the task was coin-flipping), showed a very large enhancement ( $d = 3.09$ ). This means, therefore, that the effect of completion-contingent rewards would have shown an even stronger undermining if the Kruglanski et al. study had not been designed to find an unusual circumstance where expected monetary rewards would have a positive rather than negative effect.

*Relation to previous meta-analyses.* The meta-analysis by Eisenberger and Cameron (1996) is the only one that explored this category of rewards. They reported a nonsignificant effect on both measures ( $d = -0.12$  for free-choice and  $-0.05$  for self-reports), compared with our significant effects ( $d = -0.44$  and  $-0.17$ , respectively). Their effect sizes, after outliers were removed, were

based on only 8 studies for free-choice and only 6 for self-reports, which in each case are fewer than half the numbers we used. Appendix A shows the differences in what studies from this category were included and how they were treated in our meta-analyses relative to that of Cameron and Pierce (1994).

### Performance-Contingent Rewards

Ryan et al. (1983) defined this type of reward as being given explicitly for doing well at a task or for performing up to a specified standard. Thus, in the Ryan et al. study, all participants in the performance-contingent rewards condition received \$3 for having done well at the activity. In three experiments reported by Harackiewicz et al. (1984), a group of performance-contingent rewards participants received a reward for performing better than 80% of the other participants.

However, as noted earlier, there are two important complications to this rewards category. First, there is the issue of whether or not the control group also got the feedback conveyed by the reward. Twelve performance-contingent rewards studies used only control groups that got feedback, 24 used only no-feedback control groups, and 4 (Efron, 1976; L. W. Goldstein, 1977; Harackiewicz, 1979; Ryan et al., 1983) used both.

The second issue is that although the definition of performance-contingent rewards used in many studies involved all rewarded participants getting the rewards that signify having performed well, some studies gave less than maximum rewards to some or all participants who performed (or were said to have performed) less than optimally or who were in a less-than-maximum rewards condition. These participants believed that a larger reward was possible if they performed better, so the reward conveyed less than positive (and in some cases, very negative) feedback. The procedures that conveyed explicit or implicit negative feedback existed both in studies that used feedback control groups and in studies that used no-feedback control groups.

Three studies with feedback control groups had conditions with negative feedback (Lee, 1982; Rosenfield et al., 1980; Salancik, 1975). For example, in Rosenfield et al. (1980), study participants got a small reward for performing in the bottom 15% of all participants, and the control group got the comparable negative feedback without the reward.

Seven studies with no-feedback control groups had conditions where participants got less than maximum rewards (Daniel & Esser, 1980; Dollinger & Thelen, 1978; Karniol & Ross, 1977; Luyten & Lens, 1981; Pittman et al., 1977; T. W. Smith & Pittman, 1978; Weiner & Mander, 1978). In the Daniel and Esser (1980) study, for example, rewarded participants were told that they could earn up to \$2 and were then given \$1, implying that their performance was less than optimal, perhaps at about the 50th percentile. In the Dollinger and Thelen (1978) study, children were told that they would get a reward for each good puzzle they did, and in many cases, the children did not get maximum rewards. In the Karniol and Ross (1977) study, the rewards group that was referred to as low performance was given the lesser of two rewards because it did not meet the standard. In the Luyten and Lens (1981) study, participants who got performance-contingent rewards were told that they would receive rewards for each model they solved faster than 50% of the participants. They did not get a reward for every model, implying that their performance was below the median. In

the Pittman et al. (1977) study, undergraduates played the "Gravitation" game and could earn between a nickel and a quarter on each of 10 trials, so some, if not all, received less than the maximum reward. In the T. W. Smith and Pittman (1978) study, undergraduates worked on the "Labyrinth" game and received monetary rewards that varied as a function of how well they performed on each trial. In the Weiner and Mander (1978) study, participants were told that their rewards depended on how successful they were in decoding groups of words in cartoons. They believed that, depending on their performance, their earnings would vary.

These two issues concerning control groups and feedback are examined in our meta-analyses in an attempt to achieve homogeneity of effect sizes for this diverse reward-contingency category. Specifically, we consider four categories of performance-contingency studies (or conditions within studies): those with no-feedback-control groups where rewarded participants all got maximum rewards, those with no-feedback-control groups where rewarded participants got less than maximum rewards, those with feedback-control groups where all participants got positive feedback, and those with feedback-control groups where all participants got negative feedback.

*Free-choice behavior.* The performance-contingent rewards studies are listed in Table 6. In the initial analysis of all performance-contingent rewards, we combined effects within studies if they had both types of control groups or both types of feedback. The composite effect size for the 32 studies that had a behavioral measure was  $d = -0.28$  ( $CI = -0.38, -0.18$ ), indicating that, in general, performance-contingent rewards significantly undermine free-choice intrinsic motivation.

A test of heterogeneity was significant,  $Q_w(31) = 68.06, p < .001$ , so we examined whether performance-contingent-reward effects differed for children and undergraduates and found that they did not,  $Q_b(1) = 3.37, ns$ . We then examined whether the four groups composed as a function of type of control groups and type of feedback moderated the effects of performance-contingent rewards and found that they did,  $Q_b(3) = 25.06, p < .001$ . There were 38 effect sizes included in this analysis because each of 6 studies had conditions in two groups.

Composite effects sizes were as follows. For 18 studies of maximum rewards (signifying high competence) compared with no-feedback controls, there was significant undermining,  $d = -0.15$  ( $CI = -0.31, -0.00$ ),  $p < .05$ . This set of effects was homogeneous,  $Q_w(17) = 24.39, ns$ . For 7 studies of less-than-maximum rewards (signifying less than optimal performance) compared with no-feedback controls, there was significant undermining,  $d = -0.80$  ( $CI = -1.03, -0.57$ ), but the effects were heterogeneous,  $Q_w(6) = 14.19, p < .05$ . When the Karniol and Ross (1977) study was dropped as an outlier, the set of effects became homogeneous,  $Q_w(5) = 6.95, ns$ , and the composite  $d$  was  $-0.88$  ( $CI = -1.12, -0.65$ ). For 10 studies of rewards signifying good performance compared with control groups that got the same positive feedback, there was significant undermining,  $d = -0.20$  ( $CI = -0.37, -0.03$ ), and the effects were homogeneous,  $Q_w(9) = 8.59, ns$ . Finally, for 3 studies of rewards signifying poor performance compared with control groups that got the same negative feedback, there was no effect,  $d = -0.03$  ( $CI = -0.37, 0.31$ ), and there was homogeneity,  $Q_w(2) = 3.08, ns$ . Thus, in three of the four instances, performance-contingent rewards

significantly undermined intrinsic motivation. Clearly, this approach to differentiating performance-contingent studies was appropriate and useful, because we had to exclude only one outlier to achieve homogeneity in all four of the groups.

An inspection of the composite effect sizes indicates that by far the most detrimental type of performance-contingent rewards—indeed, the most detrimental type of rewards—is one that is commonly used in applied settings, namely, one in which rewards are administered as a direct function of people's performance. If people do superlatively, they get large rewards, but if they do not display optimal performance, they get smaller rewards. Chi-square tests indicated that this type of performance-contingent rewards was more detrimental than each of the other three categories, but that the other three did not differ from each other. The comparisons of studies with no-feedback control-groups in which participants got less than maximal rewards with each of the other three categories were as follows: First, compared with studies with no-feedback control groups in which all participants got maximal rewards,  $\chi^2 = 20.80, p < .001$ ; second, compared with studies with comparable positive-feedback control groups,  $\chi^2 = 16.98, p < .001$ ; and finally, compared with studies with comparable negative-feedback control groups,  $\chi^2 = 13.54, p < .001$ .

The case of negative feedback and rewards, compared with negative feedback without rewards, is interesting because the data showed no effect for the rewards themselves. Apparently, negative feedback undermined intrinsic motivation sufficiently so that there was little left to be affected by the rewards. Thus, it appears that the combination of negative feedback and rewards is very detrimental relative to no feedback and no rewards, but it may not be more detrimental than negative feedback alone. However, as so few studies have compared negative feedback and rewards with just negative feedback, that part of the conclusion remains tentative.

*Self-reported interest.* Thirty performance-contingent rewards studies listed in Table 6 used self-reports. The composite  $d = -0.01$  ( $CI = -0.10, 0.08$ ) indicates no effect for performance-contingent rewards on interest/enjoyment. This set of effects was heterogeneous,  $Q_w(29) = 46.99, p < .05$ , so we compared the effects for children versus college students and, as in the free-choice analysis, found no difference,  $Q_b(1) = 1.45, ns$ . We then compared the four categories of effects that we had used for free-choice, but that comparison was not significant for self-reports,  $Q_b(3) = 6.46, ns$ . When we removed one outlier (Harackiewicz, 1979), the set of 29 effects became homogeneous,  $Q_w(28) = 36.04, ns$ , and the composite  $d = -0.01$  ( $CI = -0.08, 0.10$ ).

To summarize, performance-contingent rewards had a negative effect on free-choice intrinsic motivation but not on self-reported interest/enjoyment.

*Relation to previous meta-analyses.* The Tang and Hall (1995), Eisenberger and Cameron (1996), and Cameron and Pierce (1994) meta-analyses examined performance-contingent rewards (although Eisenberger and Cameron referred to them as quality-dependent). Tang and Hall concluded that performance-contingent rewards undermined intrinsic motivation, although they did not do separate analyses for the different dependent measures. Eisenberger, Cameron, and Pierce included 8 free-choice studies in the performance-contingent category after removing outliers and re-

**Table 6**  
*Effects of Performance-Contingent Rewards on Free-Choice Behavior and Self-Reported Interest, Presented as Cohen's *d*, Corrected for Sample Size*

| Study                              | Free-choice <i>d</i> | <i>n<sub>e</sub>, n<sub>c</sub></i> <sup>a</sup> | Sample <sup>b</sup> | Self-report <i>d</i> | Type <sup>c</sup> |
|------------------------------------|----------------------|--|---------------------|----------------------|-------------------|
| Bartelme (1983) D                  |                      | 99, 49   | 2                   | -0.03                | 3                 |
| Boggiano & Ruble (1979)            | -0.26                | 20, 20   | 1                   |                      | 1                 |
| Boggiano et al. (1985)             | -0.27                | 26, 13   | 1                   |                      | 1                 |
| Brewer (1980) D                    | -0.20                | 48, 24   | 1                   | 0.12                 | 1                 |
| Chung (1995)                       | -1.60                | 5, 5   | 1                   |                      | 1                 |
| D. S. Cohen (1974) D               | -0.16                | 52, 52   | 2                   | 0.12                 | 3                 |
| Dafoe (1985) D                     | 0.02                 | 72, 24   | 1                   | 0.56                 | 3                 |
| Daniel & Esser (1980)              | -0.75                | 16, 16   | 2                   | -0.71                | 2                 |
| Dollinger & Thelen (1978)          | -0.54                | 36, 12   | 1                   | 0.00 <sup>d</sup>    | 2                 |
| Efron (1976) D                     |                      | 16, 13   | 2                   | 0.50                 | 1                 |
| Efron (1976) D                     |                      | 27, 15   | 2                   | -0.57                | 3                 |
| Enzle et al. (1991)                | -0.53                | 40, 10   | 2                   |                      | 1                 |
| Fabes (1987), Exp. 1               | -0.77                | 19, 19   | 1                   |                      | 1                 |
| L. W. Goldstein (1977) D           | 0.13                 | 16, 16   | 1                   | -0.51                | 1                 |
| L. W. Goldstein (1977) D           | -0.28                | 16, 16   | 1                   | -0.43                | 3                 |
| Greene & Lepper (1974)             | -0.57                | 15, 14   | 1                   |                      | 1                 |
| Harackiewicz (1979)                |                      | 31, 16   | 1                   | -0.28                | 1                 |
| Harackiewicz (1979)                |                      | 31, 15   | 1                   | -1.47                | 3                 |
| Harackiewicz & Manderlink (1984)   |                      | 47, 47   | 1                   | 0.33                 | 3                 |
| Harackiewicz et al. (1984), Exp. 1 | 0.07                 | 32, 32   | 2                   | 0.04                 | 3                 |
| Harackiewicz et al. (1984), Exp. 2 | -0.40                | 15, 15   | 2                   | -0.16                | 3                 |
| Harackiewicz et al. (1984), Exp. 3 | 0.04                 | 26, 26   | 2                   | 0.32                 | 3                 |
| Harackiewicz et al. (1987)         |                      | 24, 24   | 1                   | -0.11                | 3                 |
| Hyman (1985) D                     | 0.07                 | 32, 32   | 1                   |                      | 1                 |
| Karniol & Ross (1977)              | 0.12                 | 10, 10   | 1                   |                      | 1                 |
| Karniol & Ross (1977)              | 0.37                 | 10, 10   | 1                   |                      | 2                 |
| Kruglanski et al. (1975), Exp. 2   |                      | 20, 20   | 1                   | -0.74                | 3                 |
| Lee (1982) D                       | -0.41                | 35, 34   | 2                   | 0.26                 | 3                 |
| Lee (1982) D                       | -0.30                | 35, 34   | 2                   | 0.39                 | 4                 |
| Luyten & Lens (1981)               | -0.82                | 10, 10   | 2                   | 0.05                 | 2                 |
| Orlick & Mosher (1978)             | -0.54                | 12, 14   | 1                   |                      | 1                 |
| Pallak et al. (1982)               | -0.23                | 15, 12   | 1                   |                      | 1                 |
| Patrick (1985) D                   | 0.00 <sup>d</sup>    | 30, 31   | 1                   | -0.48                | 1                 |
| Picek (1976) D                     | 0.00 <sup>d</sup>    | 10, 10   | 2                   | -0.69                | 1                 |
| Pittman et al. (1977)              | -1.46                | 60, 20   | 2                   | -0.20                | 2                 |
| Rosenfield et al. (1980)           | 0.13                 | 15, 14   | 2                   | -0.02                | 3                 |
| Rosenfield et al. (1980)           | 0.48                 | 15, 14   | 2                   | 0.46                 | 4                 |
| Ryan et al. (1983)                 | 0.06                 | 32, 16   | 2                   | 0.00 <sup>d</sup>    | 1                 |
| Ryan et al. (1983)                 | -0.45                | 32, 32   | 2                   | 0.00 <sup>d</sup>    | 3                 |
| Salancik (1975)                    | -0.77                | 19, 21   | 2                   | -0.39                | 3                 |
| Salancik (1975)                    | 0.09                 | 19, 18   | 2                   | 0.40                 | 4                 |
| Shiffman-Kaufman (1990) D          | 0.38                 | 20, 20   | 1                   | 0.00                 | 1                 |
| W. E. Smith (1975) D               | -0.53                | 24, 24   | 2                   | 0.00                 | 1                 |
| T. W. Smith & Pittman (1978)       | -0.94                | 66, 33   | 2                   | 0.00 <sup>d</sup>    | 2                 |
| Taub & Dollinger (1977)            |                      | 124, 124   | 2                   | 0.00 <sup>d</sup>    | 1                 |
| Tripathi & Agarwal (1988)          | 0.87                 | 20, 10   | 2                   | 1.01                 | 1                 |
| Weinberg & Jackson (1979)          |                      | 40, 40   | 2                   | 0.00 <sup>d</sup>    | 1                 |
| Weiner & Mander (1978)             | -0.59                | 30, 30   | 2                   | 0.00 <sup>d</sup>    | 2                 |

*Note.* Studies that involved multiple comparisons based on type of control group and implicit or explicit feedback have a separate listing for each comparison. D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup>The sample size for the experimental condition (*n<sub>e</sub>*) is reported first, followed by the sample size for the control group (*n<sub>c</sub>*). <sup>b</sup>In this column, 1 signifies that the participants were children, and 2 signifies that they were college students. <sup>c</sup>Type refers to whether the control group got feedback and whether the rewarded participants got the maximum reward: 1 = no feedback control, maximum reward; 2 = no feedback control, nonmaximum reward; 3 = positive feedback control; 4 = negative feedback control. <sup>d</sup>In these cases, *d* = 0.00 was assigned to the study because its authors failed to report relevant cell means.

### *Summary of Results From the Primary Meta-Analyses*

Figure 1 presents a summary of the findings from the primary meta-analysis for free-choice behavior, and Figure 2 presents a summary of the findings concerning self-reported interest.

**Free-choice behavior.** Results confirm that, for free-choice behavior, there was significant undermining by all rewards, tangible rewards, expected rewards, engagement-contingent rewards, completion-contingent rewards, and performance-contingent rewards. Only positive feedback showed enhancement, and only unexpected and task-noncontingent rewards showed no effects.

Interesting age effects also appeared, with the effects of verbal rewards being significantly less positive for children than for college students and with the effects of engagement-contingent rewards being significantly more negative for children than for college students. Although there were no significant age effects in the free-choice analyses for the other differentiated reward and reward-contingency categories, an analysis of age effects for all tangible rewards taken together did reveal a significant difference,  $Q_b(1) = 4.18, p < .04$ . For 57 tangible-reward studies of free-choice behavior with children, the composite  $d = -0.39$  ( $CI = -0.47, -0.32$ ), and for 38 with college students, the composite  $d = -0.27$  ( $CI = -0.36, -0.19$ ). Thus, it seems that tangible rewards are detrimental to the free-choice intrinsic motivation of both children and undergraduates but that they tend to be more detrimental to the intrinsic motivation of children than of college students. Furthermore, the analyses showed that verbal rewards did not enhance free-choice intrinsic motivation for children but did for college students. These age effects are particularly important in light of the strong advocacy by Cameron and Pierce (1994), Eisenberger and Cameron (1996), and others of the use of rewards in educational settings with children.

The results that emerged from the primary meta-analyses for free-choice intrinsic motivation are as predicted by the motivational and attributional approaches and are consistent with various narrative summaries (e.g., Deci & Ryan, 1985; Ryan et al., 1983). The results are also generally consistent with the results of three previous, though less extensive, meta-analyses (Rummel & Feinberg, 1988; Tang & Hall, 1995; Wiersma, 1992). Only one previous meta-analysis reported considerably weaker results, with the only significant undermining being for aggregate categories that involved engagement-contingent rewards (Eisenberger & Cameron, 1996). The Eisenberger and Cameron analysis is thus discrepant from the others. It is because of this discrepancy that we have included the appendixes showing exactly how our meta-analyses differed from those of Eisenberger, Cameron, and Pierce.

**Self-reported interest.** In general, the results for self-reported interest paralleled those for free-choice behavior, although they were substantially weaker. The analysis of overall reward effects was not significant, though the analysis of tangible rewards, expected rewards, engagement-contingent rewards, and completion-contingent rewards all showed significant undermining of intrinsic interest. The only category expected to undermine self-reported interest that did not was performance-contingent rewards. As with free-choice behavior, verbal rewards enhanced intrinsic interest. No significant age effects emerged for self-reported interest in the differentiated reward and reward-contingency categories. However, because there had been an overall age effect, with tangible rewards being significantly more undermining of children's than

of college students' free-choice behavior, we also did an analysis of age effects for all tangible rewards on self-reported interest and found that tangible rewards were significantly more detrimental for children than college students,  $Q_b(1) = 4.78, p < .05$ . Thus, although the differential effects of rewards on the self-reports of children versus undergraduates were considerably less prevalent than the differential effects on free-choice behavior, there is the suggestion that tangible rewards may also affect children's self-reports more negatively.

**Free-choice behavior and self-reported interest.** As already mentioned, the results for free-choice behavior were generally parallel to those for self-reported interest. In other words, in most cases—with the notable exception of performance-contingent rewards—enhancement, undermining, and no effects appeared for the same reward and contingency categories in each analysis, although the effects for self-reports were much weaker. To examine the relation between these two widely used measures of intrinsic motivation, we identified 17 studies that included both measures and reported the correlations between them.

The range of correlations was  $-0.20$  to  $0.71$ , with an average of  $0.35 (p < .0001)$ . Thus, the two measures are related, although not strongly. Of the 17 studies, 5 were with children and 12 with college students. The average  $r$  was quite similar for the two groups, with the average  $r$  for children being  $0.43 (p < .01)$  and for college students being  $0.32 (p < .001)$ . We return to a discussion of free-choice behavior versus self-reported interest in the discussion section.

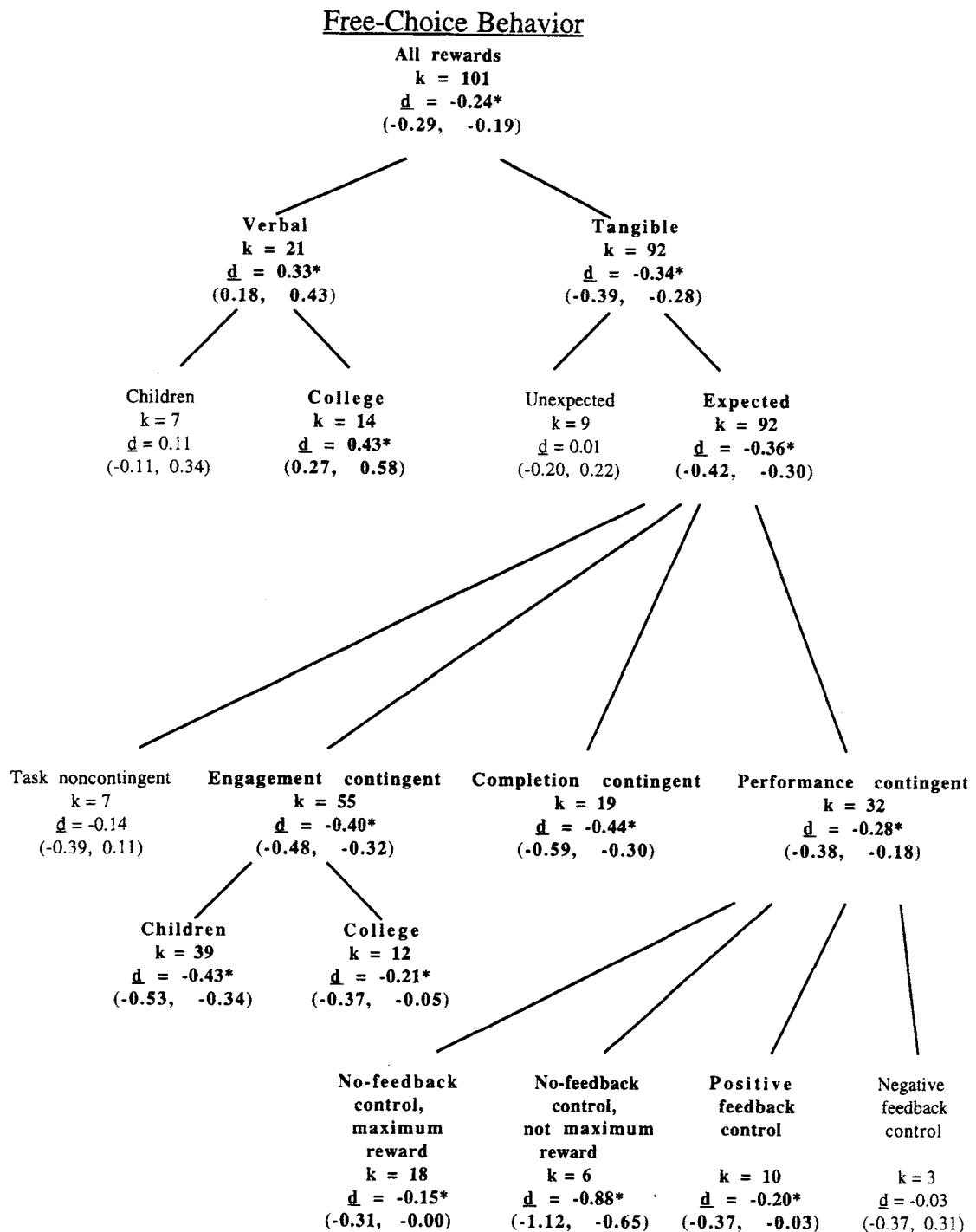
### *An Additional Reward-Contingency Category*

We have completed the report of the two primary hierarchical meta-analyses, one for free-choice behavior and the other for self-reports, but we now consider one additional category, namely task-contingent rewards. This category was included in the Ryan et al. (1983) typology, and the majority of studies examining reward effects on intrinsic motivation actually used the task-contingent terminology. As noted earlier, the separation of task-contingent rewards into completion-contingent and engagement-contingent was done very recently to examine whether these two types of task-contingent rewards have different effects on intrinsic motivation. Examination of the results of the primary meta-analyses for both free-choice (Figure 1) and self-reports (Figure 2) indicates that the effects of these two types of rewards were virtually identical (except that significant age differences appeared for the effects of engagement-contingent but not completion-contingent rewards on free-choice behavior). For the purpose of comparing our results with those of previous meta-analyses, we now present the meta-analyses for task-contingent rewards.

### *Task-Contingent Rewards*

Data for this category were obtained by combining the data from Tables 4 and 5 for the engagement-contingent and completion-contingent studies.

**Free-choice behavior.** For 74 studies with a free-choice measure, the composite effect size for task-contingent rewards showed highly significant undermining,  $d = -0.39$  ( $CI = -0.46, -0.32$ ). Of course, this set of studies was heterogeneous,  $Q_w(73) = 178.4, p < .001$ , and there was an age effect,  $Q_b(1) = 8.39, p < .01$ . For



**Figure 1.** Summary statistics for effects of types of rewards and reward-contingencies on free-choice behavior at various levels of aggregation. The symbol  $k$  refers to the number of studies (and thus the number of effect sizes) in each composite effect size. Each  $d$  entered into the composite was corrected for sample size. All reported composite effect sizes that are not further differentiated are homogeneous. The pair of numbers in parentheses represents the 95% confidence interval for the composite effect size. Boldface type and an asterisk next to a  $d$  value indicate that the composite effect size is significant. This figure includes all unpublished dissertations and all studies that had imputed  $d$  values of 0.00 because they did not report cell means.

### Self-Reported Interest

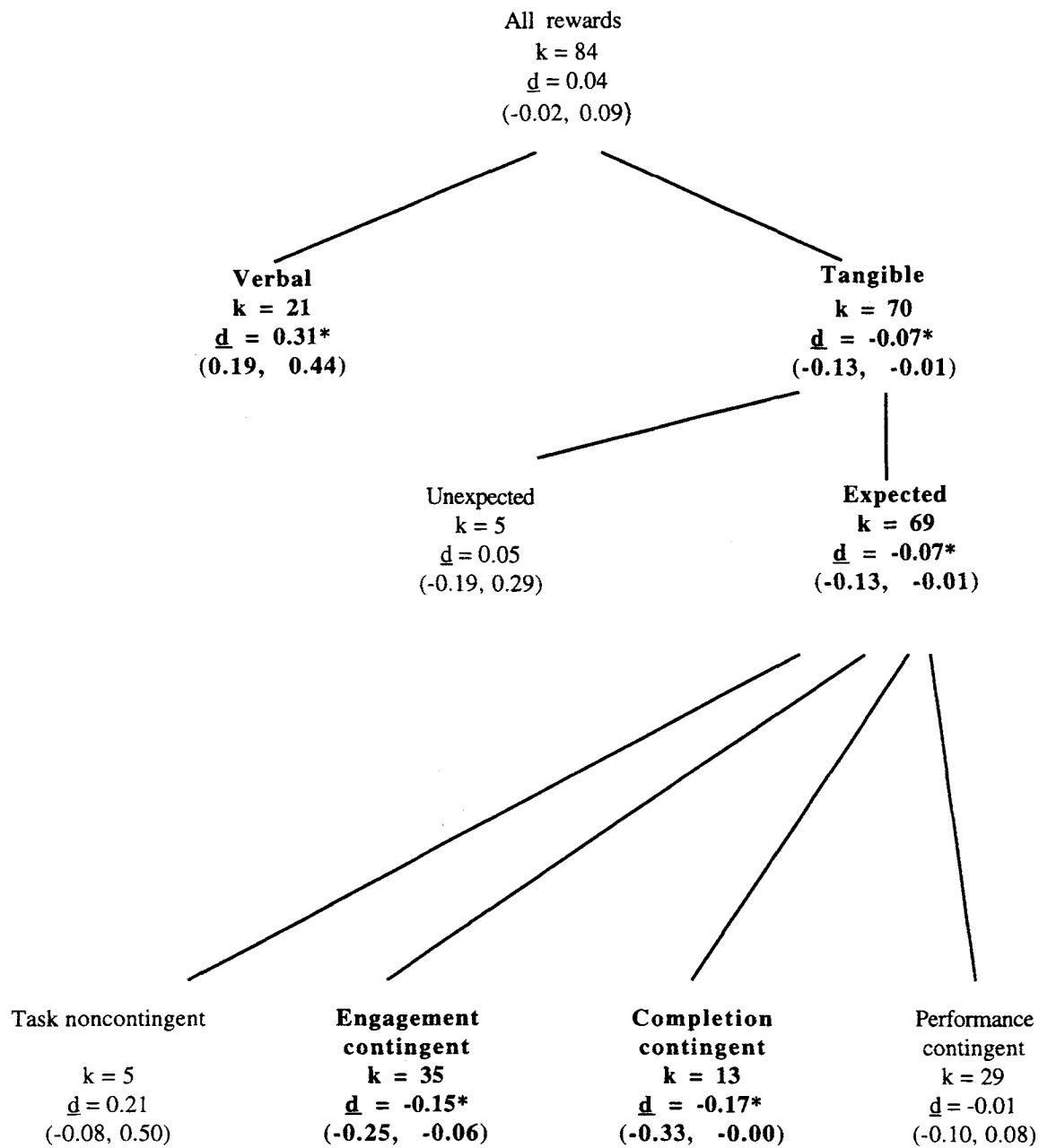


Figure 2. Summary statistics for effects of types of rewards and reward-contingencies on self-reported interest at various levels of aggregation. The symbol  $k$  refers to the number of studies (and thus the number of effect sizes) in each composite effect size. Each  $d$  entered into the composite was corrected for sample size. All reported composite effect sizes that are not further differentiated are homogeneous. The pair of numbers in parentheses represent the 95% confidence interval for the composite effect size. Boldface type and an asterisk next to a  $d$  value indicate that the composite effect size is significant. This figure includes all unpublished dissertations and all studies that had imputed  $d$  values of 0.00 because they did not report cell means.

children, the results were  $k = 51$ ;  $d = -0.46$  (CI =  $-0.54$ ,  $-0.30$ );  $Q_w(50) = 128.64$ ,  $p < .001$ . Four outliers (Boggiano et al., 1982; Danner & Lonky, 1981, Exp. 2; Morgan, 1983, Exp. 1; Swann & Pittman, 1977, Exp. 2) were removed, and the composite  $d = -0.44$  (CI =  $-0.53$ ,  $-0.35$ ). The set of effects was homogeneous,  $Q_w(46) = 60.98$ , ns. For undergraduates, the results were  $k = 23$ ;  $d = -0.25$  (CI =  $-0.36$ ,  $-0.13$ );  $Q_w(22) = 41.36$ ,  $p < .01$ . Two outliers (Brennan & Glover, 1980; Deci, 1972b) were removed, and the homogeneous effects had a composite  $d = -0.33$  (CI =  $-0.45$ ,  $-0.20$ );  $Q_w(20) = 27.36$ , ns.

*Self-reported interest.* For the 48 studies with a self-report measure of intrinsic motivation, the composite effect size for task-contingent rewards showed significant undermining,  $d = -0.12$ , (CI =  $-0.20$ ,  $-0.04$ ). This set of effect sizes was also heterogeneous,  $Q_w(47) = 84.08$ ,  $p < .001$ . With the removal of two outliers (Kruglanski et al., 1975, Exp. 1; Wimperis & Farr, 1979), the results became  $d = -0.16$  (CI =  $-0.24$ ,  $-0.07$ );  $Q_w(45) = 54.99$ , ns.

To summarize, as expected, task-contingent rewards significantly undermined both free-choice behavior and self-reported interest. We mentioned earlier that the Ryan et al. (1983) typology of reward contingencies did not break the task-contingent rewards category into the engagement-contingent and completion-contingent categories because prior research had seemed to indicate that the two types of reward-contingencies had similar effects. The present meta-analyses confirm that, showing that the effects of task-contingent, engagement-contingent, and completion-contingent rewards are all essentially the same except that there was no significant age effect for completion-contingent rewards on free-choice behavior.

*Relation to previous meta-analyses.* This finding provides a conceptual replication of the results of each of the four previous meta-analyses. Although Rummel and Feinberg (1988) did not classify studies in terms of reward contingencies, most of the studies they used involved task-contingent rewards, and although some of the studies used by Wiersma (1992) compared task-contingent rewards with no rewards and others compared task-contingent rewards with task-noncontingent rewards, the findings from both the Rummel and Feinberg and the Wiersma meta-analyses do imply an undermining of intrinsic motivation by task-contingent rewards. Tang and Hall (1995) tested whether task-contingent rewards undermined intrinsic motivation relative to no rewards and found a significant effect size of  $d = -0.51$  for 62 comparisons. Their results are comparable to ours, although, as noted earlier, their analysis included the six dependent measures.

Finally, Cameron and Pierce (1994) reported significant undermining of free-time behavior in 45 task-contingent rewards studies ( $d = -0.32$ ) before outliers were removed and in 40 studies after outliers were removed ( $d = -0.23$ ). They also reported no effect on self-reports in 21 studies before outliers were removed ( $d = -0.07$ ) and in 19 studies after outliers were removed ( $d = -0.08$ ). Eisenberger and Cameron (1996) did not report results for task-contingent rewards.

### Supplemental Analyses

We considered four important issues by performing supplemental comparisons or meta-analyses. The first concerned the relation

between unpublished doctoral dissertations and published articles. The second concerned whether the undermining of intrinsic motivation by expected tangible rewards that are contingent on performance (i.e., engagement-, completion-, and performance-contingent rewards) is simply a transitory effect.

The third important issue concerned the interest value of the target task. The field of reward effects on intrinsic motivation was begun to assess whether rewards affect people's intrinsic motivation for intrinsically interesting activities. However, some studies included uninteresting as well as interesting tasks to investigate whether the nature of the target task would moderate reward effects. We performed two supplemental analyses: First, we tested for moderation by examining reward effects for interesting tasks relative to uninteresting tasks; second, we added the uninteresting-task conditions to the data used in the primary meta-analyses and then repeated those analyses to examine whether the results would replicate with the dull tasks included.

The fourth issue addressed in a supplemental meta-analysis tested a central hypothesis of CET, namely that informationally administered rewards have a positive effect on intrinsic motivation, relative to controllingly administered rewards. We identified 4 studies that explored the informational versus controlling administration of verbal rewards and only 1 that examined tangible rewards, so we performed the supplemental meta-analysis on only the verbal rewards.

### Consideration of Published Versus Unpublished Studies

There is the possibility that well-executed studies that do not yield significant effects do not get published, in which case the published studies provide an inaccurate representation of the actual effects that have emerged from past research. To examine this possibility, we compared the effects of studies from unpublished doctoral dissertations with those of studies from published articles to see if they were different. (Doctoral dissertations that had been published were included among the published articles rather than the dissertations.) Doctoral dissertations represent an excellent database for unpublished studies because they are carefully scrutinized by supervisors and doctoral committees to ensure that they are well designed and carried out. Thus, the use of dissertations that may have remained unpublished because they did not have significant results ensures that the unpublished studies examined were not unpublished because of flawed methodology.

We compared the effects of verbal rewards on free-choice behavior and self-reported interest for published versus unpublished studies and found that they did not differ on either measure:  $Q_b(1) = 1.18$ , ns, for free-choice;  $Q_b(1) = 2.18$ , ns, for self-reports. This suggests that the published studies of verbal-reward effects may be a reasonable representation of all carefully done studies of that phenomenon. We then compared the effects of all tangible rewards on free-choice behavior and self-reported interest for published versus unpublished studies and found that they did not differ on either measure:  $Q_b(1) = 1.07$ , ns, for free-choice;  $Q_b(1) = 1.65$ , ns, for self-reports. This suggests that the published studies of tangible-reward effects may also be a reasonable representation of all well-designed studies of that phenomenon.

### *Is Undermining Transitory?*

Studies exploring reward effects on intrinsic motivation have been laboratory experiments, typically with one session, in which a reward group is compared with a no-reward group on a measure of intrinsic motivation taken immediately after the experimental phase of the single session. This paradigm has been criticized by behaviorists on two grounds (see, e.g., Davidson & Bucher, 1978; Reiss & Sushinsky, 1976; Vasta & Stirpe, 1979).

First, the immediate assessment of intrinsic motivation has led behaviorists to assert that undermining is merely a transitory effect, like temporary satiation (Eisenberger & Cameron, 1996). To investigate this, we compared studies with immediate assessments versus those with delayed assessments. Studies with delayed assessments were virtually all with children, and most included only a free-choice measure. Thus, we limited this analysis to studies of children's free-choice behavior. Thirty-two experiments had immediate assessments of free-choice intrinsic motivation, 12 had delayed assessment of free-choice done within a week of the experimental phase, and 14 had delayed assessment done more than a week after the experimental manipulation, typically within 2 weeks.

An analysis of these studies indicated that the three timings of assessment did not affect the results,  $Q_b(2) = 5.26, p < .01$ . A supplemental meta-analysis of the three groups of effects on the free-choice measure showed that for 32 studies with immediate assessments, the composite effect was  $d = -0.35$  ( $CI = -0.45, -0.25$ ), although these effects were heterogeneous,  $Q_w(31) = 69.77, p < .001$ . With the removal of 2 outliers (Boggiano et al., 1982; Swann & Pittman, 1977, Exp. 2), the effects became homogeneous,  $Q_w(29) = 42.96, ns$ , and the composite effect was  $d = -0.40$  ( $CI = -0.51, -0.29$ ). For 12 studies with delayed assessments done within a week, the composite effect was  $d = -0.49$  ( $CI = -0.68, -0.31$ ), and was homogeneous,  $Q_w(11) = 18.91, ns$ . Finally, for 14 studies with assessments done more than a week later, the composite effect was  $d = -0.55$  ( $CI = -0.71, -0.40$ ), but was heterogeneous,  $Q_w(13) = 52.33, p < .001$ . With the removal of 2 outliers (Amabile, Hennessey, & Grossman, 1986, Exp. 1; Morgan, 1983, Exp. 1), the effects became homogeneous,  $Q_w(11) = 17.90, ns$ , and the composite effect was  $d = -0.53$  ( $CI = -0.71, -0.35$ ). The homogeneous effects for this supplemental meta-analysis are presented in Table 7 and indicate quite clearly that the phenomenon of extrinsic tangible rewards undermining intrinsic motivation is not merely transitory. However, because only studies with children used delayed assessments, there is no basis for generalizing these results to adults.

Table 7  
*Effects of Expected Tangible Rewards on Free-Choice Behavior as a Function of Time of Assessment (Child Studies Only)*

| Timing        | <i>k</i> | <i>d</i> | 95% Confidence interval |
|---------------|----------|----------|-------------------------|
| Immediately   | 30       | -0.40    | -0.51, -0.29            |
| Within 1 week | 12       | -0.49    | -0.68, -0.31            |
| After 1 week  | 12       | -0.53    | -0.71, -0.35            |

*Note.* The composite *d* values shown in this table are all homogeneous. The values for the composite *d* values before outliers were removed are provided in the text.

Behaviorists have also argued that the use of a short, single-session reward period is not ecologically valid, so they have called for the use of multiple-session reward periods (see, e.g., Reiss & Sushinsky, 1975). To examine this, we searched for studies that used multiple reward sessions and found relatively few. The studies we found were done primarily with a token-economy format using participants' prereward free-choice time as the control for their postreward free-choice time. Only 2 studies met our inclusion criterion of having a no-reward control group, so we examined those 2 studies. The 2 studies, which were both included in our primary meta-analyses, were done by Mynatt et al. (1978), with 5 experimental and 5 control participants, and by Vasta and Stirpe (1979), with 4 experimental and 5 control participants. Both used a pre-post design, so we compared changes in the experimental and control groups from before to after the experimental phase. In the Mynatt et al. study, the free-choice measure showed an effect size of  $d = -0.11$  ( $CI = -1.34, 1.14$ ), and in the Vasta and Stirpe study, the free-choice measure showed an effect size of  $d = -0.14$  ( $CI = -1.09, 0.80$ ). Thus, both showed small though nonsignificant undermining, but with a total of only 9 rewarded participants in the 2 studies combined, there is no basis for drawing conclusions.

Eisenberger and Cameron (1996) discussed 4 other studies with within-subject designs (Davidson & Bucher, 1978; Feingold & Mahoney, 1975; Mawhinney, Dickinson, & Taylor, 1989; Vasta, Andrews, McLaughlin, Stirpe, & Comfort, 1978). Actually, Vasta et al. (1978) reported 2 studies although Cameron and Pierce (1994) analyzed only the first, so there were 5 relevant studies. Eisenberger and Cameron stated that none of the studies reported significant undermining and argued that these studies were representative of real-life circumstances.

It is true that none of the 5 studies showed significant undermining, but one can conclude very little. First, the 5 studies combined had a total of only 23 participants (3 for Davidson & Bucher [1978]; 5 for Feingold & Mahoney [1975]; 3 for Mawhinney et al. [1989]; and 6 in each of the Vasta et al. [1978] studies), making it tenuous to generalize the null effects. Further, given that none of these studies had control groups, one must be careful about interpretations, particularly since there appears to be substantial variability in the rates of pretreatment behavior. We do agree that studies of interesting behaviors that examine repeated administration of rewards over time, have appropriate no-reward control groups, and use reasonable sample sizes would make a valuable addition to the literature in this field.

### *Interesting Versus Boring Tasks*

Research examining tangible reward effects on intrinsic motivation has focused on interesting activities, and the results have shown that tangible rewards undermine intrinsic motivation for such activities. A few studies considered the effects of tangible rewards on both interesting and boring tasks in order to help clarify the parameters of the undermining effect. Because rewards would not be expected to undermine intrinsic motivation for an activity that is not intrinsically motivating, most investigators predicted that initial interest would moderate the effects of rewards, with rewards undermining intrinsic motivation for interesting tasks but maintaining or enhancing intrinsic motivation for boring tasks. As already noted, in spite of the moderator results found in several

**Table 8**  
*Effects of Tangible Rewards on Intrinsic Motivation, Presented as Cohen's *d*, Corrected for Sample Size as a Function of Interest Level of Task*

| Study                   | Dependent variable <sup>a</sup> | Interesting task <i>d</i> <sup>b</sup> | Boring task <i>d</i> <sup>b</sup> | Contingency <sup>c</sup> |
|-------------------------|---------------------------------|--|-----------------------------------|--------------------------|
| Calder & Staw (1975)    | S                               | -0.76 (10, 10)                         | 1.01 (10, 10)                     | CC                       |
| Chung (1995)            | F                               | -1.43 (10, 10)                         | 1.69 (10, 10)                     | EC, PC                   |
| Daniel & Esser (1980)   | F                               | -0.74 (16, 16)                         | -0.28 (16, 16)                    | PC                       |
| Daniel & Esser (1980)   | S                               | -0.71 (16, 16)                         | 0.04 (16, 16)                     | PC                       |
| Eisenstein (1985)       | F                               | -0.50 (18, 10)                         | 0.29 (16, 6)                      | CC                       |
| Griffith (1984) D       | F                               | -0.73 (42, 42)                         | -0.31 (42, 42)                    | EC                       |
| Hamner & Foster (1975)  | S                               | -0.20 (35, 15)                         | 0.11 (35, 15)                     | EC, CC                   |
| Hitt et al. (1992)      | F                               | -0.67 (30, 15)                         | 0.46 (30, 15)                     | EC                       |
| Hitt et al. (1992)      | S                               | -0.48 (30, 15)                         | -0.15 (30, 15)                    | EC                       |
| Loveland & Olley (1979) | F                               | -1.17 (6, 6)                           | 1.17 (6, 6)                       | EC                       |
| McLoyd (1979)           | F                               | -1.01 (18, 9)                          | 0.55 (18, 9)                      | CC                       |
| Mynatt et al. (1978)    | F                               | -0.11 (5, 5)                           | 1.10 (5, 5)                       | EC                       |
| Newman & Layton (1984)  | F                               | -0.36 (21, 10)                         | 0.40 (21, 10)                     | EC                       |
| A. T. Smith (1980) D    | F                               | -0.75 (21, 27)                         | 0.08 (21, 27)                     | EC                       |
| Wilson (1978) D         | F                               | -0.34 (23, 23)                         | 0.00 (23, 23) <sup>d</sup>        | EC                       |
| Wilson (1978) D         | S                               | -0.06 (23, 23)                         | 0.00 (23, 23) <sup>d</sup>        | EC                       |

*Note.* D after the name and date of a study indicates that it is an unpublished doctoral dissertation.

<sup>a</sup> In this column, F denotes a free-choice dependent measure, and S denotes self-reported interest. <sup>b</sup> Numbers in parentheses after the *d* value refer to sample sizes, with the experimental conditions reported first, followed by the sample size for the control groups. <sup>c</sup> Contingency refers to the following: EC = engagement-contingent; CC = completion-contingent; PC = performance-contingent. <sup>d</sup> In these cases, *d* = 0.00 was assigned to the study because relevant cell means were not reported.

studies, Cameron and Pierce (1994) and Eisenberger and Cameron (1996) collapsed across interesting and boring task conditions in their meta-analyses, implying that from a behaviorist perspective, initial task interest has no theoretical meaning.

We found 13 studies that experimentally manipulated interest level of the task, with one task defined as interesting and having initial interest ratings above the midpoint on the interest scale and another task defined as uninteresting and having low initial interest ratings. Six of these studies had been included in the meta-analyses by Eisenberger, Cameron, and Pierce with the interesting and uninteresting conditions combined. In this supplemental meta-analysis of tasks, we begin with those 13 studies, examining the effects of rewards on the interesting versus uninteresting tasks within the same study. Table 8 presents the effect sizes of rewards on interesting versus dull tasks for the 13 studies. Eight of the studies used only a free-choice measure, 2 used only a self-report measure, and 3 used both.

For the 11 studies with a free-choice measure, reward effects on interesting activities differed significantly from reward effects on uninteresting activities,  $Q_b(1) = 31.67, p < .0001$ . The composite effect for interesting tasks showed significant undermining,  $d = -0.68$  ( $CI = -0.89, -0.47$ ), and for uninteresting tasks did not show a significant effect,  $d = 0.18$  ( $CI = -0.03, 0.39$ ). For the 5 studies with a self-report measure, reward effects on interesting activities also differed significantly from reward effects on uninteresting activities,  $Q_b(1) = 4.88, p < .05$ . The composite effect for interesting tasks was  $d = -0.37$  ( $CI = -0.67, -0.07$ ), thus showing significant undermining for tangible rewards, but for uninteresting tasks there was not a significant effect,  $d = 0.10$  ( $CI = -0.09, 0.40$ ). Thus, it is clear that the effects of tangible rewards are different when the tasks are interesting versus dull and that the reliable undermining of intrinsic motivation by tangible rewards does not extend to dull-boring tasks.

#### Overall Meta-Analyses With Dull Task Included

As noted earlier, one reason that Eisenberger, Cameron, and Pierce's results were weaker than ours is that they included dull tasks. To examine how much the inclusion of dull tasks would affect our significant results, we repeated the primary meta-analyses after including the dull-task data for the 13 relevant studies that had experimentally manipulated interest value of the task.

We first present the results of the meta-analysis for free-choice data. Because no dull-task data were included in the verbal rewards, unexpected rewards, or task-noncontingent rewards categories, the statistics remain the same for them and are not repeated below. With data from dull-task conditions included, the free-choice results for the aggregate categories were as follows: For all rewards,  $k = 101$ ;  $d = -0.21$  ( $CI = -0.27, -0.16$ );  $Q_w(100) = 274.74, p < .0001$ ; for all tangible rewards,  $k = 92$ ;  $d = -0.30$  ( $CI = -0.36, -0.24$ );  $Q_w(91) = 217.65, p < .0001$ ; and for all expected rewards,  $k = 92$ ;  $d = -0.32$  ( $CI = -0.38, -0.26$ );  $Q_w(91) = 217.42, p < .0001$ . For the differentiated contingency categories, if a category was heterogeneous, we removed outliers to make it homogeneous so it would be comparable to the results reported in other meta-analyses. For engagement-contingent rewards,  $k = 55$ ;  $d = -0.35$  ( $CI = -0.45, -0.27$ );  $Q_w(54) = 143.73, p < .001$ , so with 5 outliers removed to achieve homogeneity,<sup>5</sup>  $k = 50$ ;  $d =$

<sup>5</sup> The outliers for engagement-contingent free-choice were Boggiano et al. (1982), Brennan and Glover (1980), Danner and Lonky (1981, Exp. 2), Morgan (1983, Exp. 1), and Swann and Pittman (1977, Exp. 2). The outlier for completion-contingent free-choice was Deci (1972b). The outlier for performance-contingent free-choice was Pittman et al. (1977). The outliers for completion-contingent self-reports were Kruglanski et al. (1975, Exp. 1) and Wimperis and Farr (1979). The outlier for performance-contingent self-reports was Harackiewicz (1979).

$-0.33$  (CI =  $-0.41, -0.25$ );  $Q_w(49) = 65.99$ , ns; for completion-contingent rewards,  $k = 20$ ;  $d = -0.32$  (CI =  $-0.46, -0.19$ );  $Q_w(19) = 33.05$ ,  $p < .05$ , so with 1 outlier removed to achieve homogeneity,  $k = 19$ ;  $d = -0.40$  (CI =  $-0.54, -0.25$ );  $Q_w(18) = 23.16$ , ns; and for performance-contingent rewards,  $k = 32$ ;  $d = -0.27$  (CI =  $-0.37, -0.17$ );  $Q_w(31) = 64.39$ ,  $p < .001$ , so with 1 outlier removed to achieve homogeneity,  $k = 31$ ;  $d = -0.23$  (CI =  $-0.33, -0.13$ );  $Q_w(30) = 45.43$ , ns. To summarize, in the analysis of free-choice data, although the inclusion of the dull tasks weakened the composite  $d$  values slightly, there were no instances in which the inclusion made a composite effect nonsignificant that had been significant.

For self-reports, with data from dull-task conditions included, the results for the aggregate categories were as follows: For all rewards,  $k = 84$ ;  $d = 0.04$  (CI =  $-0.01, 0.10$ );  $Q_w(83) = 198.22$ ,  $p < .001$ ; for all tangible rewards,  $k = 70$ ;  $d = -0.06$  (CI =  $-0.12, 0.00$ ),  $p < .05$ ;  $Q_w(69) = 122.17$ ,  $p < .0001$ ; and for all expected rewards,  $k = 69$ ;  $d = -0.06$  (CI =  $-0.12, -0.00$ ),  $p < .05$ ;  $Q_w(68) = 116.38$ ,  $p < .0001$ . As with the free-choice data, in the contingency categories, we present composite effect sizes for homogeneous sets of effects. For engagement-contingent rewards,  $k = 35$ ;  $d = -0.15$  (CI =  $-0.24, -0.06$ );  $Q_w(34) = 37.57$ , ns; for completion-contingent rewards,  $k = 15$ ;  $d = -0.02$  (CI =  $-0.13, 0.17$ );  $Q_w(14) = 41.00$ ,  $p < .001$ , so with 2 outliers removed to achieve homogeneity,  $k = 13$ ;  $d = -0.10$  (CI =  $-0.26, 0.05$ );  $Q_w(12) = 16.30$ , ns; and for performance-contingent rewards,  $k = 30$ ;  $d = -0.01$  (CI =  $-0.10, 0.08$ );  $Q_w(29) = 44.97$ ,  $p < .05$ , so with 1 outlier removed to achieve homogeneity,  $k = 29$ ;  $d = 0.01$  (CI =  $-0.07, -0.10$ );  $Q_w(28) = 34.00$ , ns. To summarize, in the analysis of self-report data, the inclusion of the dull tasks weakened the composite  $d$  values slightly, and there was one instance in which the inclusion made a composite effect size drop from significance to nonsignificance, namely, the completion-contingent rewards category. There was no evidence of enhancement in any category (except, of course, verbal rewards). All other aggregate and differentiated categories were not affected in terms of whether they were significant. Thus, the inclusion of the dull-task conditions did affect the results, but it accounted for a relatively small amount of the discrepancy between the results of our meta-analyses and those of Eisenberger, Cameron, and Pierce.

### Informational Versus Controlling Verbal Rewards

CET predicts that, in general, positive feedback enhances intrinsic motivation, and indeed, the meta-analyses provided strong confirmation of this prediction. However, within the theory, there is an important caveat distinguishing between positive feedback that is administered informationally and positive feedback that is administered controllingly. CET predicts that controllingly administered positive feedback leads to significantly less intrinsic motivation or interest than does informationally administered positive feedback. As shown in Table 9, 4 studies, 1 using a self-report measure (Kast & Connor, 1988) and the other 3 using a free-choice measure (Pittman et al., 1980; Ryan, 1982; Ryan et al., 1983), included both types of positive feedback and found substantially different results for the two types of verbal rewards.

The composite effect size for the 4 studies of informationally versus controllingly administered positive feedback was  $d = -0.78$  (CI =  $-1.02, -0.54$ ), indicating a large and highly significant effect in which controllingly administered positive feedback decreased intrinsic motivation relative to informationally administered positive feedback. Three of the 4 studies also had no-feedback control groups, and a very important result emerged from these studies. The composite effect size comparing informational positive feedback with a no-feedback control group yielded a composite  $d = 0.66$  (CI =  $0.28, 1.03$ ), whereas the effect size comparing controlling positive feedback with a no-feedback group yielded a composite  $d = -0.44$  (CI =  $-0.82, -0.07$ ). This suggests that whereas informationally administered positive feedback enhances intrinsic motivation, controllingly administered positive feedback may actually decrease it.

In these studies, the positive feedback was made controlling by saying, for example, that someone had done very well, "just as you should" (Ryan et al., 1983, p. 745, italics added). The should statement is what made the feedback controlling and thus undermining of intrinsic motivation. Given how easy it is to use such statements in the real world, this finding emphasizes the practical complexity of providing positive feedback without its being detrimental.

One study (Ryan et al., 1983) examined the effects of informationally versus controllingly administered performance-contingent tangible rewards as well as the effects of informationally versus

Table 9  
Effects of Informational Versus Controlling Verbal Rewards on Intrinsic Motivation,  
Presented as Cohen's  $d$ , Corrected for Sample Size

| Study                                  | Type of comparison                            |   |   |
|--|---|---|---|
|  | Informational vs.<br>controlling <sup>a</sup> | Neutral vs.<br>informational <sup>a</sup> | Neutral vs.<br>controlling <sup>a</sup> |
| Kast & Connor (1988)                   | -1.39 (30, 30) <sup>b</sup>                   | 0.43 (30, 30)                             | -1.04 (30, 30)                          |
| Pittman et al. (1980)                  | -0.65 (36, 36)                                | 1.06 (12, 12)                             | 0.00 (12, 12)                           |
| Ryan (1982)                            | -0.65 (64, 64)                                |   |   |
| Ryan et al. (1983)                     | -0.63 (16, 16)                                | 0.81 (16, 16)                             | 0.21 (16, 16)                           |
| Overall: $d$ (95% confidence interval) | -0.78 (0.54, 1.02)                            | 0.66 (0.28, 1.03)                         | -0.44 (-0.82, -0.07)                    |

<sup>a</sup> In the first and third columns, a negative  $d$  signifies undermining of intrinsic motivation in the controlling condition relative to the informational (first column) or neutral (third column) conditions. In the second column, a positive  $d$  signifies enhancement of intrinsic motivation in the informational condition relative to the neutral condition. <sup>b</sup> Ns for the conditions in each comparison are listed in parentheses after the effect sizes in the rows signifying particular studies.

controllingly administered verbal rewards. CET also predicts a significant difference for informational versus controlling tangible rewards. The results for the verbal-rewards comparisons in the Ryan et al. study are presented in Table 9, and the results for the informationally versus controllingly administered tangible-rewards comparisons in that study are as follows. With cell sizes of 16, the effect size for the informational versus controlling comparison was  $d = -0.70$  ( $CI = -1.42, 0.01$ ), suggesting that, as with verbal rewards, controllingly administered tangible rewards undermined intrinsic motivation relative to informationally administered tangible rewards. The effect size comparing informationally administered tangible rewards with no rewards yielded a  $d = 0.44$  ( $CI = -0.26, 1.14$ ), suggesting enhancement by informational rewards, whereas the effect size comparing controllingly administered tangible rewards with no rewards yielded a  $d = -0.29$  ( $CI = -0.98, 0.41$ ), suggesting undermining by controlling rewards. The pattern and magnitude of these three  $d$  values closely parallel the pattern and magnitude for informational versus controlling positive feedback and thus provide additional support for CET.

## Discussion

The picture that emerged from these meta-analyses of 128 well-controlled experiments exploring the effects of extrinsic rewards on intrinsic motivation is clear and consistent. In general, tangible rewards had a significant negative effect on intrinsic motivation for interesting tasks, and this effect showed up with participants ranging from preschool to college, with interesting activities ranging from word games to construction puzzles, and with various rewards ranging from dollar bills to marshmallows. The only exception was that, although performance-contingent rewards undermined free-choice behavior, they did not affect self-reported interest. On the other hand, verbal rewards—or what is usually labeled positive feedback in the motivation literature—had a significant positive effect on intrinsic motivation, although the effect on free-choice behavior was found for college students but not children. In accord with the motivational and attributional theories, which predict that tangible rewards do not undermine intrinsic motivation when they are not expected or not contingent on task behavior, there was no significant undermining in those cases. The undermining effects were all considerably stronger for the free-choice measure than for the self-report measure, but the results for self-reports did parallel those for free-choice, with the already noted exception of performance-contingent rewards.

### *Fit of the Theories to the Data*

#### *Cognitive Evaluation Theory*

CET has made differentiated predictions about the effects of rewards on intrinsic motivation for interesting activities, and the results of the present meta-analyses strongly support the theory's predictions. Because the theory expects verbal and tangible rewards to work in opposite directions, it makes no predictions about the effects of all reward effects when verbal and tangible rewards are combined. The significant undermining of free-choice behavior by all rewards and the nonsignificant effect for self-reported in-

terest have little theoretical meaning. They are simply a function of the number of tangible-reward studies versus verbal-reward studies that were identified and the relative effect size of each of those categories on the two dependent measures.

As expected, because positive feedback affirms competence if it is not administered controllingly, it enhanced intrinsic motivation. Also as predicted, when positive feedback was administered controllingly, it diminished intrinsic motivation. There was, however, the finding not predicted by CET that positive feedback did not enhance the free-choice intrinsic motivation of children. CET's approach to this issue would be to investigate why children apparently experienced the feedback as more controlling than did college students. That investigation remains to be done, although we provide speculations later in this section.

Concerning expected tangible rewards that explicitly require involvement with the target task, CET predicts that they typically are experienced as controlling and thus decrease intrinsic motivation. As predicted, the results showed that both engagement-contingent and completion-contingent rewards did decrease free-choice intrinsic motivation and self-reported interest. Further, performance-contingent rewards, which are the most complex category, decreased free-choice behavior as expected, with the undermining tending to be less than for engagement-contingent and completion-contingent rewards, but they did not affect self-reports. CET analyzes performance-contingent rewards in terms of the informational and controlling aspects of rewards working in opposite directions, and it points out that the informational aspect of performance-contingent rewards is stronger than that of any other tangible-reward category. Thus, the theory expects that, in general, the undermining is less for performance-contingent rewards than for the other two reward contingencies. The fact that this reward contingency did not affect self-reported interest even though it negatively affected free-choice behavior is addressed as a methodological issue later in the article. Further, as predicted, unexpected rewards and task-noncontingent rewards did not have a detrimental effect because participants were not doing the task in order to get those rewards, so they did not feel controlled by them. In sum, then, the pattern of effects in our meta-analyses provided very strong support for CET in that the effects were virtually all as predicted by the theory.

Relatively little work has directly tested mediational effects of perceived locus of causality and perceived competence on people's intrinsic motivation for interesting activities. Some studies have found that positive feedback enhanced intrinsic motivation as mediated by perceived competence (see, e.g., Blanck, Reis, & Jackson, 1984; Sansone, 1989; Vallerand & Reid, 1984), and some studies of performance-contingent rewards have shown that competence valuation is an important mediator of intrinsic motivation (Harackiewicz, Abrahams, & Wageman, 1987) and serves to make the informational aspect of rewards more salient (Harackiewicz et al., 1984). Less well researched has been the concept of perceived self-determination (i.e., perceived locus of causality) as a mediator of reward effects. Reeve and Deci (1996) showed that perceived self-determination as well as perceived competence mediated the effects of competition and competitive outcomes on intrinsic motivation, but a similar analysis has not been done with respect to

reward effects.<sup>6</sup> In sum, the present meta-analysis provides strong support for CET in the sense that the predictions made by the theory are supported by the data, but additional work remains to be done to explain the age effects and to show that perceived self-determination and perceived competence mediate these predicted effects.

### *The Attributional Approach*

This approach, typified in the overjustification work of Lepper (1981), focuses on the discounting of intrinsic motivation in the presence of contingencies and the enhancement of intrinsic motivation in the presence of competence cues and makes predictions that are quite similar to those of CET. Results of the current meta-analyses are generally consistent with this approach also. As with CET, the attributional approach has not been well examined in terms of specific mediational processes, but a few studies have shown that rewarded subjects attribute their task performance less to internal causes (see, e.g., Brockner & Vasta, 1981; Pittman et al., 1977). Additional studies involving more formal mediational analyses would be useful.

### *The Behavioral Approaches*

These theories received very little support from the current meta-analyses. The most differentiated and inclusive of the behavioral approaches (Eisenberger & Cameron, 1996) predicted that tangible rewards undermine intrinsic motivation only in what that study called the performance-independent conditions (i.e., the engagement-contingent and task-noncontingent conditions) and enhance or leave unchanged intrinsic motivation in any conditions where rewards are dependent on performance. The theoretical reasoning was that because performance-independent rewards convey that people cannot control desired outcomes, the rewards thus produce helplessness. This helplessness effect, Eisenberger and Cameron argued, has often been misinterpreted as the undermining of intrinsic motivation.

There are several problems with this reasoning. First, past evidence showing helplessness following receipt of uncontrollable positive outcomes is very scant, so predicting that engagement-contingent or task-noncontingent rewards induce helplessness does not have an established empirical basis. Further, the engagement-contingent category does not fit the definition of helplessness-inducing situations. Engagement-contingent rewards are explicitly given for merely doing a task. For example, in a study by Morgan (1981), subjects were told that they could select a reward "in return for working on the puzzle for 10 minutes" (p. 813). Thus, with engagement-contingent rewards, control over the rewards is clear and requires behavior that participants can perform. They control the desired outcome (viz., getting the reward) precisely by doing the target behavior. We have found nothing in the writings on helplessness theory by Seligman (1975) and his colleagues suggesting that such a situation induces helplessness, and we maintain that only task-noncontingent rewards, which do not require any explicit task behavior, could possibly fit the definition of helplessness-inducing conditions.

The present meta-analyses fail to support Eisenberger and Cameron's (1996) theoretical approach. First, using both dependent measures, our analyses showed that engagement-contingent re-

wards significantly undermined intrinsic motivation whereas task-noncontingent rewards did not. Because, as already noted, helplessness theory predicts that task-noncontingent rewards are more likely to induce helplessness than engagement-contingent rewards, the results, which showed just the opposite, fail to support that model. In fact, an examination of the results in the two articles by the Eisenberger, Cameron, and Pierce group reveals that they too found undermining for engagement-contingent but not for task-noncontingent rewards, so their own analyses argue against their model. Further, contrary to the Eisenberger and Cameron predictions, the present meta-analyses show that completion-contingent rewards undermined both free-choice intrinsic motivation and self-reported interest, and performance-contingent rewards undermined free-choice intrinsic motivation. Performance-contingent rewards did not affect self-reported interest, but that is the condition where Eisenberger and Cameron predicted the strongest enhancement.

Because the Eisenberger and Cameron (1996) helplessness hypothesis was formulated post hoc following their meta-analyses, it has received little direct test, so we looked for studies that might provide such a test and found two. In Weiner (1980), one group of rewarded participants was told that success at an anagram task was largely a matter of being lucky enough to find the right combinations, and the other group of rewarded participants was told that success depended on ability. Rewards for solving a puzzle due to luck should, if any rewards do, produce helplessness but in fact led to a large enhancement of free-choice behavior ( $d = 1.19$ ). In contrast, rewards for solving a puzzle due to ability should, according to the Eisenberger and Cameron model, lead to enhance-

---

<sup>6</sup> A study by Boal and Cummings (1981) has been portrayed by Cameron and Pierce (1994) and others as a process study that disconfirmed CET, although in fact it did not provide a test of the theory at all. First, as already noted in Footnote 1, there was not a no-payment control group; instead, all groups got an hourly wage to work on a boring task, and the so-called rewards groups got an unexpected raise. Thus, because the raise was initially unexpected and was subsequently given task-noncontingently (i.e., was paid as part of an hourly wage that was not dependent on performance), it would not be predicted by CET to decrease intrinsic motivation even if the task had been interesting and there had been an appropriate no-reward control group. Furthermore, Boal and Cummings claimed to have assessed feelings of competence and self-determination, and they reported that that variable did not mediate reward effects on intrinsic motivation. However, their assessments were wholly inappropriate. Specifically, they had a 5-item assessment labeled "feelings of competence and self-determination," two examples of which are: "Doing my job well increases my feelings of self-esteem," and "I always work as hard as I can" (p. 301). Consider the first item, which we suppose they intended to capture feelings of self-determination, although self-esteem does not necessarily vary with self-determination (Deci & Ryan, 1995). The item is worded in what is essentially an if-then format: If I did my job well, then I would feel self-esteem. It does not assess either whether participants felt that they had done their job well or whether they had high self-esteem. The second item assesses whether they always expend effort and has nothing to do with whether they feel either competent or self-determined. The other three items all have similar problems. Studies by Freedman and Phillips (1985) and Overskeid and Svartdal (1998) also claimed to have tested the processes of CET, but both studies used dull-boring tasks and are thus not relevant to this discussion. As noted earlier, the processes of internalization proposed by Deci and Ryan (1991) as being pertinent to motivation for uninteresting activities are different from the processes described by CET concerning motivation for interesting activities.

ment yet here led to undermining ( $d = -0.69$ ). Further, in a study by Kruglanski et al. (1975, Exp. 1), one group got rewards for correctly guessing coin-flip outcomes (a task that is clearly luck based), and the rewards led to substantial enhancement of self-reported interest ( $d = 3.09$ ), whereas another group got rewards for correctly constructing models (an effort/ability task), and the rewards led to undermining ( $d = -0.87$ ). Thus, both studies yielded strong results that were exactly opposite of what the Eisenberger and Cameron model predicts.

Eisenberger and Cameron (1998) stated that they had done another meta-analysis in which they compared performance-contingent rewards that had specific performance standards (e.g., rewards for surpassing the 80th percentile) with comparable-positive-feedback control groups, indicating that there were 4 such studies with the free-choice measure and 7 with the self-report measure and reporting that in both cases, these performance-contingent rewards significantly enhanced intrinsic motivation. They did not provide any methodological details, so it is difficult to know what comparisons they made, but there are two important points to keep in mind. First, there is the issue of whether the feedback control group was told before beginning that they would be evaluated. In the Harackiewicz et al. (1984, Exp. 1) study, which was of this type, performance-contingent rewards did not enhance either measure of intrinsic motivation relative to a no-reward group that got comparable feedback, but performance-contingent rewards did enhance both measures relative to a group that was told they would be evaluated and then got positive feedback. Thus, the more appropriate interpretation is that the rewards did not enhance intrinsic motivation but that the evaluation undermined it. Second, in some studies of performance-contingent rewards with performance standards, there were conditions where the standards signified good performance and conditions where they signified poor performance. One such study was Rosenfield et al. (1980), in which Cameron and Pierce collapsed across positive and negative performance conditions. In fact, as shown in Table 6 of our results, the rewards signifying good performance did not enhance intrinsic motivation relative to a control group with comparable positive feedback, but the rewards signifying poor performance did enhance intrinsic motivation relative to a control group with comparable negative feedback.

To evaluate the Eisenberger and Cameron (1998) claim, we examined the studies of performance-contingent rewards that met the criteria of specific performance standards and positive-feedback control groups. These were a subset of the positive-feedback control-group studies in Table 6 with the code of 3 in the final column. Six studies with a free-choice measure met the criteria (Harackiewicz et al., 1984, Exps. 1, 2, & 3; Lee, 1982; Rosenfield et al., 1980; Salancik, 1975). We used only conditions for which the normative feedback indicated that performance was above the 50th percentile, so the negative-feedback conditions in Rosenfield et al. and Salancik were excluded. For these six homogeneous effects,  $d = -0.21$  (CI =  $-0.44, 0.03$ ),  $Q_w(5) = 6.61, ns$ , thus contradicting the Eisenberger and Cameron claim that this type of reward enhances free-choice intrinsic motivation. Ten studies with a self-report measure met the criteria (the same 6 plus Efron, 1976; Harackiewicz, 1979; Harackiewicz & Manderlink, 1984; Harackiewicz et al., 1987). However, for 2 studies (Harackiewicz, 1979; Harackiewicz et al., 1987), not all performance-

contingent rewards groups were relevant because not all worked with the normative standards, so the  $d$  values used in this analysis for those 2 studies differed somewhat from the  $d$  values reported in Table 6. For these 10 homogeneous effect sizes,  $d = -0.02$  (CI =  $-0.20, 0.17$ ),  $Q_w(9) = 16.43, ns$ , also contradicting the Eisenberger and Cameron claim that this type of reward enhances self-reported interest. In sum, there is no support for the Eisenberger and Cameron claim that was said to have come from their 1998 meta-analysis.

Further, concerning the adequacy of the behavioral approaches for explaining intrinsic motivation phenomena, it is worth noting that, unlike the motivational and attributional approaches, the behavioral theories have not been generalized to conceptually similar issues such as how other external events such as deadlines (Amabile et al., 1976), competition (Deci, Betley, et al., 1981; Reeve & Deci, 1996), and choice (Zuckerman et al., 1978) affect intrinsic motivation. Finally, behavioral theories have not distinguished between interesting and uninteresting tasks but have claimed that reward effects on all tasks should be the same even though the data show clearly that they are not.

### *The Free-Choice and Self-Report Measures*

Results of the two primary meta-analyses indicate that the pattern of reward effects on intrinsic motivation was similar when assessed with the two commonly used dependent measures, although the effects on free choice were much stronger than on self-reports. It is thus interesting to consider further the two measures (Quattrone, 1985) that were found in the current analyses to have an average correlation of about 0.35 for those studies where the correlations were reported.

Each has been useful as a measure of intrinsic motivation in some studies, but each has a disadvantage that could cause different potential problems. With self-reports of interest/enjoyment for a rewarded task, people may confuse their enjoyment of the reward with their interest in the task. To the extent that this happens, it inflates self-reports of interest in the rewards conditions relative to the control group, leading to an underestimation of the undermining by rewards. As for the free-choice measure, the potential problem is that the extrinsic motivation manipulated in the experimental phase could possibly persist into the free-choice period, leading to some free-choice behavior that is a reflection of extrinsic, rather than intrinsic, motivation. To the extent that this happens, it inflates the free-choice measure in the rewards group relative to the control group, leading to an underestimation of the undermining by tangible rewards.

Elsewhere (Ryan, Koestner, & Deci, 1991), we have argued that the best way to ensure one is assessing intrinsic motivation is to measure both free-choice behavior and self-reported interest and to consider them intrinsic motivation only when they correlate within conditions or studies. Unfortunately, relatively few studies have measured both and reported the correlations, so we are left with the question of which measure is the more accurate reflection of intrinsic motivation in rewards studies. We believe that free-choice behavior is a more valid measure of intrinsic motivation and thus that the set of findings for the free-choice measure is the more accurate indicator of the actual effects of rewards on intrinsic motivation. There are several reasons for this.

First, the free-choice behavioral measure is typically unobtrusive; in other words, participants typically believe that the experimenter is not going to know whether they have persisted at the activity during the free-choice period. Thus, demand characteristics and other interpersonal considerations are not likely to affect the free-choice measure. On the other hand, participants know that the experimenter is going to see their answers on the self-report measure, and the questions are typically quite transparent, so demand characteristics could play a larger role with that measure. Second, the reliability of the self-report measures of interest is unknown, and because many studies have used only one (or very few) items, the measures may suffer from low reliability. Third, the primary problem associated with the free-choice measure (viz., that extrinsic motivation could underlie free-choice persistence) is not likely to operate when the extrinsic motivation is the pursuit of a tangible reward because tangible-reward contingencies can be clearly terminated. It is in conditions of ego-involved extrinsic motivation (which is an intrapsychic state that cannot be easily terminated by the experimenter) where this problem has been observed (Ryan et al., 1991). Furthermore, even if this problem had been operative in reward conditions, it would have decreased (rather than increased) the negative effect sizes for the free-choice measure. On the other hand, the primary problem associated with the self-report measure (viz., that it could reflect people's enjoyment of the reward as well as the task) could easily have been operative in these studies, leading to higher levels of self-reported interest in the rewards conditions than would otherwise have been the case. This provides a plausible account of why the undermining was weaker for the self-report than for the free-choice measure. Finally, the simple fact that we had to impute effects of 0.00 for the self-report measure in 26 out of 121 cases (21%), whereas we had to impute effects of 0.00 for the free-choice measure in only 8 out of 150 cases (5%), is itself indication that the self-report measure is less effective than the free-choice measure.

### *Age Effects*

Our meta-analyses at the level of all tangible rewards showed that tangible rewards are more detrimental for children than for college students on both measures and that verbal rewards do not enhance free-choice behavior for children although they do for undergraduates.<sup>7</sup> This set of findings has never been predicted before, and the issue has not been previously examined, so we can only speculate about what might be occurring. There are of course many differences between the capacities and experiences of children and college students that could be involved in this difference. First, it is likely that children ongoingly experience more attempts by adults to regulate their behavior with rewards and other potentially controlling contingencies than do college students. Second, college students have greater cognitive capacity for separating the informational and controlling aspects of rewards and are also more accustomed to operating with performance-goal orientations, so they may be more ready to interpret rewards as indicators of their effective performance than as controllers of their behavior (see, e.g., Elliot & Harackiewicz, 1994; Harackiewicz & Elliot, 1993). Third, undergraduates typically have had more experience as participants in research and may have more expectations about whether rewards are appropriate in such situations than do chil-

dren, and those expectations could affect how they experience the rewards.

### *Implications for Use of Rewards in the Real World*

Our findings raise deeply important questions about when and how to use rewards in real-world settings. First, we address tangible rewards, then verbal rewards.

#### *Tangible Rewards*

The findings indicate that tangible rewards offered for engaging in, completing, or doing well at a task were deleterious to intrinsic motivation, particularly as assessed with free-choice behavior, but that unexpected and task-noncontingent rewards had no effect on intrinsic motivation. Thus, if people use tangible rewards, it is necessary that they be extremely careful if they are concerned about the intrinsic motivation and task persistence of the people they are rewarding. Consider several points in this regard.

First, unexpected rewards, which were not detrimental, may be a way, upon occasion, to convey appreciation for a task well done without harming intrinsic motivation. But one must be mindful that people may begin to expect the "unexpected" rewards if they are given very often.

Second, task-noncontingent rewards were also found not to be detrimental, but these rewards, by definition, do not provide a means of expressing appreciation for doing a task or doing it well because they are not linked to the task. Thus, their potential use is quite limited. In fact, their primary utility may be in situations where rewards are necessary but are not used to convey information or to motivate performance (e.g., paying people a salary for occupying a job).

Third, the style of administering tangible rewards also influences their effect. According to CET, rewards given informationally tend to have a less negative (or more positive) effect on intrinsic motivation than do those given controlling, although only the Ryan et al. (1983) study has examined and confirmed this prediction for tangible rewards. Theirs is an important finding because it means that even though the controlling aspect of tangible rewards is very strong, it may be possible to administer tangible rewards informationally so as not to have a negative effect.

Research by Deci, Nezlek, et al. (1981), Koestner, Ryan, Bernieri, and Holt (1984), and Deci et al. (1994) suggests that making rewards more informational requires (a) minimizing the use of authoritarian style and pressuring locution, (b) acknowledging good performance but not using rewards to try to strengthen or control the behavior, (c) providing choice about how to do the

<sup>7</sup> Because tangible rewards were very detrimental for children, but the types of tangible rewards were quite varied, we did an additional pair of comparisons between the effects of tangible concrete rewards that children desire, such as toys or candy, and the effects of tangible symbolic rewards, such as a "good player" certificate. There were no differences. For free-choice, with 28 concrete-reward studies,  $d = -0.44$  ( $CI = -0.54, -0.33$ ), and with 15 symbolic-reward studies,  $d = -0.42$  ( $CI = -0.57, -0.27$ ),  $Q_b(1) = 0.03, ns$ . For self-report, with 17 concrete-reward studies,  $d = -0.24$  ( $CI = -0.36, -0.11$ ), and with 4 symbolic-reward studies,  $d = -0.18$  ( $CI = -0.44, 0.07$ ),  $Q_b(1) = 0.14, ns$ .

tasks, and (d) emphasizing the interesting or challenging aspects of the tasks. Indeed, the tangible rewards themselves are really incidental to the process of enhancing intrinsic motivation; it is these other factors, such as choice (Zuckerman et al., 1978) and non-controlling positive feedback (Ryan, 1982), that help make rewards informational and lead to enhanced intrinsic motivation.

Fourth, the fact that tangible rewards seem to be more detrimental for children than for college students has many potential implications. Clearly, it implies that such rewards must be used extraordinarily carefully with children in schools, homes, and elsewhere so as not to negatively affect their intrinsic motivation.

Finally, in the real world, when rewards are used to signify competence (i.e., when they are performance-contingent), there likely is a substantial portion of people who receive less than maximal rewards because they do not perform up to the specified standards. For these people, the failure to get the maximum reward may be experienced as negative feedback and could be highly detrimental because the reward structure is not only controlling but also conveys negative-competence information. In fact, the meta-analyses demonstrate that, relative to a no-rewards no-feedback control group, people in a performance-contingent rewards group who got less than the maximum rewards showed larger undermining of free-choice intrinsic motivation ( $d = -0.88$ ) than did people in any other reward or reward-contingency group we examined. Further, a study by Pritchard, Campbell, and Campbell (1977), which was not used in our meta-analyses because participants who did not do well did not get rewards, showed very large undermining on both free-choice intrinsic motivation ( $d = -0.95$ ) and self-reported interest ( $d = -0.76$ ).

Recently, behaviorally oriented theorists such as Eisenberger and Cameron (1996) have advocated widespread use of performance-contingent and completion-contingent rewards. From the present meta-analyses, one can see that this advocacy is inconsistent with the empirical results indicating clearly that, for interesting activities, performance-contingent rewards have a detrimental effect on free-choice persistence and that completion-contingent rewards have a detrimental effect on both measures of intrinsic motivation. Thus, the use of such rewards in schools could be very detrimental to intrinsic motivation, which has been shown to relate positively to both cognitive outcomes (see, e.g., Grolnick & Ryan, 1987; Utman, 1997), well-being (see, e.g., Deci, Nezlek, et al., 1981; Ryan & Grolnick, 1986), and behavioral persistence (Vallerand, Fortier, & Guay, 1997).

### *Verbal Rewards*

From the time of the first intrinsic motivation experiment (Deci, 1971), there has been evidence that recipients experience verbal rewards (i.e., positive feedback) differently from tangible rewards. The results of our meta-analyses confirm this and suggest that positive feedback can have an enhancing effect on intrinsic motivation. In part, this is because verbal rewards are typically unexpected. In fact, it is not exactly clear how one defines an expected verbal reward. One possibility is that verbal rewards could be praise statements that people come to expect from repeated experiences. If people always get verbal rewards when doing a particular task or when with a particular person, they might well come to expect such rewards. If people engage in interesting activities specifically because they expect to get verbal rewards, we predict

that the rewards undermine intrinsic motivation, although we know of no research testing this. The other possibility is that expected verbal rewards would be the positive feedback statements that follow task engagement when people have been told they would be evaluated. In fact, 3 studies (Harackiewicz et al., 1984, Exps. 1 and 3; W. E. Smith, 1975) have investigated the effects on free-choice intrinsic motivation of people being told they would get performance feedback and then being given positive feedback after they finished the task. We examined the effects of these conditions meta-analytically and found that anticipated evaluation (when followed by positive feedback) led to significant undermining of free-choice behavior with a composite effect size of  $d = -0.40$  (CI =  $-0.71, -0.09$ ). It seems, therefore, that expected verbal rewards may undermine intrinsic motivation just as expected tangible rewards do.

To summarize, verbal rewards may have positive effects in part because they are typically unexpected; however, according to CET, it is primarily because verbal rewards provide an affirmation of competence that they can enhance intrinsic motivation.

It is important, however, to consider the conclusion of enhancement by verbal rewards in light of three other findings. First, our meta-analysis indicated that verbal rewards did not enhance children's behavioral display of intrinsic motivation. Second, 4 studies considered in a supplemental meta-analysis indicated that the informational versus controlling administration of positive feedback moderates the effects of the feedback. When positive feedback is administered controllingly, the negative effect of the control counteracts the positive effects of the information, leading to an undermining of intrinsic motivation by the positive feedback (Kast & Connor, 1988; Pittman et al., 1980; Ryan et al., 1983). Third, although we did not investigate this meta-analytically, studies by Deci, Cascio, and Krusell (1975), Koestner, Zuckerman, and Koestner (1987), and Kast and Connor (1988) indicate that there is a tendency for verbal rewards to be experienced more controllingly by female participants than by male participants, which has led to decrements in intrinsic motivation for female participants even when there are increments for male participants. Thus, not only is it likely that verbal rewards undermine intrinsic motivation when administered controllingly (Kast & Connor, 1988; Ryan, 1982), but it is possible that they also undermine the intrinsic motivation of female participants even when they are relatively ambiguous with respect to information versus control (Deci et al., 1975; Kast & Connor, 1988).

The absence of an effect for children, the moderating effect by the interpersonal style of the rewarder, and the possible gender effects all suggest that Cameron and Pierce (1994) and Eisenberger and Cameron (1996) were premature in advocating the widespread use of verbal rewards in educational settings.

### *The Power of Rewards*

There is no lack of agreement between our viewpoint and that of the operant and neo-operant theorists about the power of rewards to control behavior. It is clear that rewards can be used as a technique of control; indeed, CET specifically proposes that it is because people are controlled by rewards that they become less intrinsically motivated. Any lack of agreement concerns the unintended consequences of rewards being used to control behavior.

It is worth noting that the rewards issue is merely a special case of a more general issue, namely, the control versus self-determination of human behavior in social contexts. Many aspects of the interpersonal environment in addition to rewards have been shown to affect intrinsic motivation—either undermining or enhancing it—as a function of being experienced primarily as controlling versus informational. Events such as deadlines (Amabile et al., 1976), evaluations (W. E. Smith, 1975), and imposed goals (Mossholder, 1980) have been found to undermine intrinsic motivation, whereas other events such as the provision of choice (Zuckerman et al., 1978) and the acknowledgment of feelings (Koestner et al., 1984) have been found to enhance intrinsic motivation. Still other events such as limits (Koestner et al., 1984) and competition (Reeve & Deci, 1996) have been found either to undermine intrinsic motivation if they are made controlling (for example, by strongly emphasizing the importance of beating the other player) or to enhance intrinsic motivation if they are made informational (for example, by focusing on the individual's playing well in a challenging situation). Similarly, studies have found classroom climates and work climates that are experienced as controlling to be associated with decrements in intrinsic motivation and related variables, relative to climates that are experienced as informational (deCharms, 1976; Deci et al., 1989; Deci, Schwartz, et al., 1981; Ryan & Grolnick, 1986).

Serious consideration of the control versus self-determination of behavior not only yields important insights into such issues as conceptual learning (Grolnick & Ryan, 1987), work performance (Baard, Deci, & Ryan, 1998), and creativity (Amabile, 1996) but also leads one to think about how to promote versus forestall the development of self-regulation (Deci & Ryan, 1991; Lepper, 1983; Ryan, 1993). Although aspects of the social environment that tend to be controlling can be effective in producing behaviors, they are quite ineffective in promoting self-regulation. Considerable research has now shown, for example, that minimizing control in the social context is important not only for maintaining intrinsic motivation but also for promoting more integrated and volitional extrinsically motivated behavior (Deci et al., 1994; Grolnick & Ryan, 1989; Ryan & Connell, 1989; G. C. Williams, Grow, Freedman, Ryan, & Deci, 1996).

### *The Meaning of Intrinsic Motivation*

In the typical intrinsic-motivation experimental paradigm, the undermining effect shows up as a postreward response rate that is below the response rate of a no-rewards comparison group (and thus, implicitly, below the preredward baseline). Because this finding is contrary to a central tenet of operant theory, it has led to controversy. From our perspective, though, the findings are not primarily about response rates. They are about an important, though somewhat vulnerable, source of human motivation that we refer to as intrinsic motivation, and they are about the broader issues of human autonomy and volition. The findings from the present meta-analyses can be understood in terms of their relevance to the issue of human autonomy only if they are thoughtfully considered with respect to the real meaning of intrinsic motivation.

Intrinsic motivation energizes and sustains activities through the spontaneous satisfactions inherent in effective volitional action. It is manifest in behaviors such as play, exploration, and challenge seeking that people often do for no external rewards. It is thus a

prototypic instance of human freedom or autonomy in that people engage in such activity with a full sense of willingness and volition. According to CET, intrinsic motivation is an inherent motivational tendency that has evolved because it entails many adaptive advantages (Ryan, Kuhl, & Deci, 1997), but it still requires environmental supports. CET proposes that the necessary supports are opportunities to satisfy the innate needs for competence and self-determination. By taking account of these inherent needs and their relation to intrinsic motivation, CET has been effective in predicting the impact of extrinsic rewards and other environmental factors on intrinsic motivation and related dependent variables.

### *The Importance of Cognitive Interpretations*

For more than 25 years, we have argued that predictions about the effects of rewards necessitate a differentiated analysis of how the rewards are likely to be interpreted by the recipients based on a consideration of the type of rewards (Deci, 1971, 1975), the type of contingency (Ryan et al., 1983), the type of participants (Deci et al., 1975), and the type of interpersonal climate within which the rewards are administered (Deci, Nezlek, et al., 1981; Ryan et al., 1983).

Each of these factors provides researchers with useful information that allows meaningful predictions about whether rewards are experienced primarily as controllers of behavior or primarily as affirmations of competence. Using this approach, we have not only been successful in predicting the pervasive undermining of intrinsic motivation by most tangible reward contingencies, but we have also been able to predict when tangible rewards are not detrimental (viz., when they are unexpected, task-noncontingent, or administered informationally) and when verbal rewards enhance versus diminish intrinsic motivation.

One of the reasons that the undermining of intrinsic motivation by tangible rewards has become what Eisenberger and Cameron (1996) called the "conventional view" (p. 1153) is that it fits well with modern cognitively oriented theories of psychology. As early as 1961, Festinger (1961) suggested that it is not the reward per se, but rather its meaning to the recipient, that determines the reward's effects. Similarly, several other social-cognitive theoretical approaches such as self-attribution (Lepper et al., 1973), motivational systems (Ford, 1992), goal orientations (Ames, 1984; Dweck, 1986; Nicholls, 1984), reactance (Brehm, 1966), activity engagement (Higgins & Trope, 1990), action identification (Vallacher & Wegener, 1987), and personal causation (deCharms, 1968) all predict that rewards, under some nontrivial conditions, undermine intrinsic motivation. Even though that phenomenon is not a central concern of each theory, each theory does emphasize that cognitive interpretations of external events affect motivation, affect, and behavior. In line with those theories, the current meta-analysis supports the importance of cognitive mediators of rewards and feedback, further emphasizing the importance of human experience in understanding how environmental events influence subsequent behavior.

### *Concluding Comments*

Careful consideration of reward effects reported in 128 experiments leads to the conclusion that tangible rewards tend to have a

substantially negative effect on intrinsic motivation, with the limiting conditions we have specified. Even when tangible rewards are offered as indicators of good performance, they typically decrease intrinsic motivation for interesting activities.

Although rewards can control people's behavior—indeed, that is presumably why they are so widely advocated—the primary negative effect of rewards is that they tend to forestall self-regulation. In other words, reward contingencies undermine people's taking responsibility for motivating or regulating themselves. When institutions—families, schools, businesses, and athletic teams, for example—focus on the short term and opt for controlling people's behavior, they may be having a substantially negative long-term effect. Furthermore, as noted by Kohn (1993), when organizations opt for the use of rewards to control behavior, the rewards are likely to be accompanied by greater surveillance, evaluation, and competition, all of which have also been found to undermine intrinsic motivation (Deci & Ryan, 1985).

Research has shown the value of being intrinsically motivated in many applied settings such as education, sports, and work environments. In addition, research on intrinsic motivation has focused attention on the more general benefits of supports for autonomy and competence for motivated persistence, performance, and well-being. Many social institutions face problems including alienation, detachment, and disengagement that could be at least partially ameliorated by promoting higher levels of intrinsic motivation and self-determination. Strategies focused on optimizing the psychological need satisfactions associated with active engagement of various tasks within specific performance settings thus offer important alternatives to the use of rewards and other social controls to motivate behavior. As research has shown, there are conditions under which tangible rewards do not necessarily undermine intrinsic motivation, but the evidence indicates clearly that strategies that focus primarily on the use of extrinsic rewards do, indeed, run a serious risk of diminishing rather than promoting intrinsic motivation.

## References

- References marked with an asterisk indicate studies included in the meta-analyses.
- Amabile, T. M. (1996). *Creativity in context*. New York: Westview Press.
- Amabile, T. M., DeJong, W., & Lepper, M. R. (1976). Effects of externally imposed deadlines on subsequent intrinsic motivation. *Journal of Personality and Social Psychology*, 34, 92–98.
- \*Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influence on creativity: The effects of contracted-for rewards. *Journal of Personality and Social Psychology*, 50, 14–23.
- Ames, C. (1984). Competitive, cooperative, and individualistic goal structures: A cognitive-motivational analysis. In R. Ames & C. Ames (Eds.) *Research on motivation: Vol. 1. Student motivation*. New York: Academic Press.
- \*Anderson, R., Manoogian, S. T., & Reznick, J. S. (1976). The undermining and enhancing of intrinsic motivation in preschool children. *Journal of Personality and Social Psychology*, 34, 915–922.
- \*Anderson, S., & Rodin, J. (1989). Is bad news always bad? Cue and feedback effects on intrinsic motivation. *Journal of Applied Social Psychology*, 19, 449–467.
- \*Arkes, H. R. (1979). Competence and the overjustification effect. *Motivation and Emotion*, 3, 143–150.
- \*Arnold, H. J. (1976). Effects of performance feedback and extrinsic reward upon high intrinsic motivation. *Organizational Behavior and Human Performance*, 17, 275–288.
- \*Arnold, H. J. (1985). Task performance, perceived competence, and attributed causes of performance as determinants of intrinsic motivation. *Academy of Management Journal*, 28, 876–888.
- Baard, P. P., Deci, E. L., & Ryan, R. M. (1998). *Intrinsic need satisfaction: A motivational basis of performance and well-being in two work settings*. Unpublished manuscript, Fordham University.
- Bakeman, R., & Gottman, J. M. (1989). *Observing interaction: An introduction to sequential analysis*. Cambridge, England: Cambridge University Press.
- \*Bartelme, L. A. (1983). *The effects of choice and rewards on intrinsic motivation and performance*. Unpublished doctoral dissertation, University of Iowa.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 2–62). New York: Academic Press.
- Bernstein, D. J. (1990). Of carrots and sticks: A review of Deci and Ryan's "Intrinsic motivation and self-determination in human behavior." *Journal of the Experimental Analysis of Behavior*, 54, 323–332.
- \*Blanck, P. D., Reis, H. T., & Jackson, L. (1984). The effects of verbal reinforcements on intrinsic motivation for sex-linked tasks. *Sex Roles*, 10, 369–387.
- Boal, K. B., & Cummings, L. L. (1981). Cognitive evaluation theory: An experimental test of processes and outcomes. *Organizational Behavior and Human Performance*, 28, 289–310.
- \*Boggiano, A. K., Harackiewicz, J. M., Bessette, J. M., & Main, D. S. (1985). Increasing children's interest through performance-contingent reward. *Social Cognition*, 3, 400–411.
- Boggiano, A. K., & Hertel, P. T. (1983). Bonuses and bribes: Mood effects in memory. *Social Cognition*, 2, 49–62.
- \*Boggiano, A. K., & Ruble, D. N. (1979). Competence and the overjustification effect. *Journal of Personality and Social Psychology*, 37, 1462–1468.
- \*Boggiano, A. K., Ruble, D. N., & Pittman, T. S. (1982). The mastery hypothesis and the overjustification effect. *Social Cognition*, 1, 38–49.
- Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.
- \*Brennan, T. P., & Glover, J. A. (1980). An examination of the effect of extrinsic reinforcers on intrinsically motivated behavior: Experimental and theoretical. *Social Behavior and Personality*, 8, 27–32.
- \*Brewer, J. (1980). *The undermining and enhancing effects of intrinsic motivation: A developmental study*. Unpublished doctoral dissertation, Ohio State University.
- \*Brockner, J., & Vasta, R. (1981). Do causal attributions mediate the effects of extrinsic rewards on intrinsic interest? *Journal of Research in Personality*, 15, 201–209.
- \*Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79, 474–482.
- \*Calder, B. J., & Staw, B. M. (1975). Self-perception of intrinsic and extrinsic motivation. *Journal of Personality and Social Psychology*, 31, 599–605.
- Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64, 363–423.
- Carton, J. S. (1996). The differential effects of tangible rewards and praise on intrinsic motivation: A comparison of cognitive evaluation theory and operant theory. *Behavior Analyst*, 19, 237–255.
- \*Chung, K. T. (1995). The effects of extrinsic reinforcement on intrinsic motivation amongst mildly mentally handicapped children. *Curriculum Forum*, 4(2), 98–114.
- \*Cohen, D. S. (1974). *The effects of task choice, monetary, and verbal*

- reward on intrinsic motivation: A closer look at Deci's cognitive evaluation theory.* Unpublished dissertation, Ohio State University.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 10*, 37-46.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- \*Crino, M. D., & White, M. C. (1982). Feedback effects in intrinsic/extrinsic reward paradigms. *Journal of Management, 8*, 95-108.
- \*Dafoe, J. L. (1985). *Use of rewards in teaching a skill: Effects on competence, self-efficacy and intrinsic motivation.* Unpublished doctoral dissertation, Stanford University.
- \*Daniel, T. L., & Esser, J. K. (1980). Intrinsic motivation as influenced by rewards, task interest, and task structure. *Journal of Applied Psychology, 65*, 566-573.
- \*Danner, F. W., & Lonky, E. (1981). A cognitive-developmental approach to the effects of rewards on intrinsic motivation. *Child Development, 52*, 1043-1052.
- Davidson, P., & Bucher, B. (1978). Intrinsic interest and extrinsic reward: The effects of a continuing token program on continuing nonconstrained preference. *Behavior Therapy, 9*, 222-234.
- deCharms, R. (1968). *Personal causation: The internal affective determinants of behavior.* New York: Academic Press.
- deCharms, R. (1976). *Enhancing motivation: Change in the classroom.* New York: Irvington.
- \*Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*, 105-115.
- \*Deci, E. L. (1972a). Effects of contingent and non-contingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance, 8*, 217-229.
- \*Deci, E. L. (1972b). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology, 22*, 113-120.
- Deci, E. L. (1975). *Intrinsic motivation.* New York: Plenum.
- Deci, E. L., Betley, G., Kahle, J., Abrams, L., & Porac, J. (1981). When trying to win: Competition and intrinsic motivation. *Personality and Social Psychology Bulletin, 7*, 79-83.
- Deci, E. L., & Cascio, W. F. (1972, April). *Changes in intrinsic motivation as a function of negative feedback and threats.* Paper presented at the meeting of the Eastern Psychological Association, Boston.
- \*Deci, E. L., Cascio, W. F., & Krusell, J. (1975). Cognitive evaluation theory and some comments on the Calder and Staw critique. *Journal of Personality and Social Psychology, 31*, 81-85.
- Deci, E. L., Connell, J. P., & Ryan, R. M. (1989). Self-determination in a work organization. *Journal of Applied Psychology, 74*, 580-590.
- Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*, 119-142.
- Deci, E. L., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology, 40*, 1-10.
- Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 39-80). New York: Academic Press.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.
- Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self-integration in personality. In R. Dienstbier (Ed.), *Nebraska Symposium on Motivation: Vol. 38. Perspectives on motivation* (pp. 237-288). Lincoln, NE: University of Nebraska Press.
- Deci, E. L., & Ryan, R. M. (1995). Human autonomy: The basis for true self-esteem. In M. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 31-49). New York: Plenum.
- Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology, 73*, 642-650.
- \*DeLoach, L. L., Griffith, K., & LaBarba, R. C. (1983). The relationship of group context and intelligence to the overjustification effect. *Bulletin of the Psychonomic Society, 21*, 291-293.
- Dickinson, A. M. (1989). The detrimental effects of extrinsic reinforcement on "intrinsic motivation." *Behavior Analyst, 12*, 1-15.
- \*Dimitroff, G. (1984). *Depression of intrinsically motivated performance by rewards: The role of frustration-mediated contrast effects.* Unpublished doctoral dissertation, University of Toronto.
- \*Dollinger, S. J., & Thelen, M. H. (1978). Overjustification and children's intrinsic motivation: Comparative effects of four rewards. *Journal of Personality and Social Psychology, 36*, 1259-1269.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.
- \*Earn, B. M. (1982). Intrinsic motivation as a function of extrinsic financial rewards and subjects' locus of control. *Journal of Personality, 50*, 360-373.
- \*Efron, B. (1976). *Effects of self-mediated competency feedback and external incentives on intrinsic motivation and quality of task performance.* Unpublished doctoral dissertation, University of Pittsburgh.
- Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist, 51*, 1153-1166.
- Eisenberger, R., & Cameron, J. (1998). Reward, intrinsic interest, and creativity: New findings. *American Psychologist, 53*, 676-679.
- \*Eisenstein, N. (1985). Effects of contractual, endogenous, or unexpected rewards on high and low interest preschoolers. *Psychological Record, 35*, 29-39.
- Elliot, A. J., & Harackiewicz, J. M. (1994). Goal setting, achievement orientation, and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 66*, 968-980.
- \*Enzle, M. E., Roggeveen, J. P., & Look, S. C. (1991). Self- versus other-reward administration and intrinsic motivation. *Journal of Experimental Social Psychology, 27*, 468-479.
- \*Fabes, R. A. (1987). Effects of reward contexts on young children's task interest. *Journal of Psychology, 121*, 5-19.
- \*Fabes, R. A., Eisenberg, N., Fultz, J., & Miller, P. (1988). Reward, affect and young children's motivational orientation. *Motivation and Emotion, 12*, 155-169.
- \*Fabes, R. A., Fultz, J., Eisenberg, N., May-Plumlee, T., & Christopher, F. S. (1989). Effects of rewards on children's prosocial motivation: A socialization study. *Developmental Psychology, 25*, 509-515.
- \*Fabes, R. A., McCullers, J. C., & Hom, H. (1986). Children's task interest and performance: Immediate vs. subsequent effects of rewards. *Personality and Social Psychology Bulletin, 12*, 17-30.
- Farr, J. L. (1976). Task characteristics, reward contingency, and intrinsic motivation. *Organizational Behavior and Human Performance, 16*, 294-307.
- \*Feehan, G. G., & Enzle, M. E. (1991). Subjective control over rewards: Effects of perceived choice of reward schedule on intrinsic motivation and behavior maintenance. *Perceptual and Motor Skills, 72*, 995-1006.
- Feingold, B. D., & Mahoney, M. J. (1975). Reinforcement effects on intrinsic interest: Undermining the overjustification hypothesis. *Behavior Therapy, 6*, 357-377.
- Festinger, L. (1961). The psychological effects of insufficient rewards. *American Psychologist, 16*, 1-12.
- Flora, S. R. (1990). Undermining intrinsic interest from the standpoint of a behaviorist. *Psychological Record, 40*, 323-346.
- Ford, M. E. (1992). *Motivating humans: Goals, emotions, and personal agency beliefs.* Newbury Park, CA: Sage.
- Freedman, S. M., & Phillips, J. S. (1985). The effects of situational performance constraints on intrinsic motivation and satisfaction: The role of perceived competence and self-determination. *Organizational Behavior and Human Decision Processes, 35*, 397-416.

- \*Goldstein, G. S. (1980). *The effects of competition and external rewards on intrinsic motivation*. Unpublished doctoral dissertation, University of New Hampshire.
- \*Goldstein, L. W. (1977). *Intrinsic motivation: The role of reward and feedback on quality of performance and subsequent interest in photography*. Unpublished doctoral dissertation, Cornell University.
- \*Greene, D., & Lepper, M. R. (1974). Effects of extrinsic rewards on children's subsequent intrinsic interest. *Child Development*, 45, 1141-1145.
- \*Griffith, K. M. (1984). *The effects of group versus individual context, initial interest, and reward on intrinsic motivation*. Unpublished doctoral dissertation, University of South Florida.
- \*Griffith, K. M., DeLoach, L. L., & LaBarba, R. C. (1984). The effects of rewarder familiarity and differential reward preference on intrinsic motivation. *Bulletin of the Psychonomic Society*, 22, 313-316.
- Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, 52, 890-898.
- Grolnick, W. S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology*, 81, 143-154.
- \*Hamner, W. C., & Foster, L. W. (1975). Are intrinsic and extrinsic rewards additive: A test of Deci's cognitive evaluation theory of task motivation. *Organizational Behavior and Human Performance*, 14, 398-415.
- \*Harackiewicz, J. M. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, 37, 1352-1363.
- \*Harackiewicz, J. M., Abrahams, S., & Wageman, R. (1987). Performance evaluation and intrinsic motivation: The effects of evaluative focus, rewards, and achievement orientation. *Journal of Personality and Social Psychology*, 53, 1015-1023.
- Harackiewicz, J. M., & Elliot, A. J. (1993). Achievement goals and intrinsic motivation. *Journal of Personality and Social Psychology*, 65, 904-915.
- \*Harackiewicz, J. M., & Manderlink, G. (1984). A process analysis of the effects of performance-contingent rewards on intrinsic motivation. *Journal of Experimental Social Psychology*, 20, 531-551.
- \*Harackiewicz, J., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: The effects of evaluation on intrinsic interest. *Journal of Personality and Social Psychology*, 47, 287-300.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Higgins, E. T., & Trope, Y. (1990). Activity engagement theory: Implications of multiply identifiable input for intrinsic motivation. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 229-264). New York: Guilford.
- \*Hitt, D. D., Marriot, R. G., & Esser, J. K. (1992). Effects of delayed rewards and task interest on intrinsic motivation. *Basic and Applied Social Psychology*, 13(4), 405-414.
- Hom, H. L. (1987). A methodological note: Time of participation effects on intrinsic motivation. *Personality and Social Psychology Bulletin*, 13, 210-215.
- \*Hyman, C. (1985). *Reward contingency, standards, and intrinsic motivation*. Unpublished doctoral dissertation, City University of New York.
- Johnson, B. T. (1993). *DSTAT 1.10: Software for the meta-analytic review of literatures* [Computer software and manual]. Hillsdale, NJ: Erlbaum.
- Johnson, B. T., & Eagly, A. H. (in press). Quantitative synthesis of social psychological research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social psychology*. New York: Cambridge University Press.
- \*Karniol, R., & Ross, M. (1977). The effect of performance-relevant and performance-irrelevant rewards on children's intrinsic motivation. *Child Development*, 48, 482-487.
- \*Kast, A., & Connor, K. (1988). Sex and age differences in response to informational and controlling feedback. *Personality and Social Psychology Bulletin*, 14, 514-523.
- Koestner, R., Ryan, R. M., Bernieri, F., & Holt, K. (1984). Setting limits on children's behavior: The differential effects of controlling versus informational styles on intrinsic motivation and creativity. *Journal of Personality*, 52, 233-248.
- \*Koestner, R., Zuckerman, M., & Koestner, J. (1987). Praise, involvement, and intrinsic motivation. *Journal of Personality and Social Psychology*, 53, 383-390.
- Kohn, A. (1993). *Punished by rewards*. Boston: Houghton-Mifflin.
- Kohn, A. (1996). By all available means: Cameron and Pierce's defense of extrinsic motivators. *Review of Educational Research*, 66, 1-4.
- Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review*, 82, 387-406.
- \*Kruglanski, A. W., Alon, S., & Lewis, T. (1972). Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology*, 8, 493-501.
- \*Kruglanski, A. W., Friedman, I., & Zeevi, G. (1971). The effects of extrinsic incentive on some qualitative aspects of task performance. *Journal of Personality*, 39, 606-617.
- \*Kruglanski, A. W., Riter, A., Amitai, A., Margolin, B., Shabtai, L., & Zaksh, D. (1975). Can money enhance intrinsic motivation? A test of the content consequences hypothesis. *Journal of Personality and Social Psychology*, 31, 744-750.
- \*Lee, J. (1982). *Effects of absolute and normative feedback and reward on performance and affect*. Unpublished doctoral dissertation, University of Iowa.
- Lepper, M. R. (1981). Intrinsic and extrinsic motivation in children: Detrimental effects of superfluous social controls. In W. A. Collins (Ed.), *Aspects of the development of competence: The Minnesota Symposium on Child Psychology* (Vol. 14, pp. 155-214). Hillsdale, NJ: Erlbaum.
- Lepper, M. R. (1983). Social-control processes and the internalization of social values: An attributional perspective. In E. T. Higgins, D. N. Ruble, & W. W. Hartup (Eds.), *Social cognition and social development* (pp. 294-330). New York: Cambridge University Press.
- \*Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28, 129-137.
- Lepper, M. R., Keavney, M., & Drake, M. (1996). Intrinsic motivation and extrinsic rewards: A commentary on Cameron and Pierce's meta-analysis. *Review of Educational Research*, 66, 5-32.
- \*Lepper, M. R., Sagotsky, G., Dafoe, J. L., & Greene, D. (1982). Consequences of superfluous social constraints: Effects on young children's social inferences and subsequent intrinsic interest. *Journal of Personality and Social Psychology*, 42, 51-65.
- \*Liberty, H. J. (1986). *Intrinsic motivation, extraversion, impulsivity, and reward in a computer game setting*. Unpublished doctoral dissertation, City University of New York.
- \*Loveland, K. K., & Olley, J. G. (1979). The effect of external reward on interest and quality of task performance in children of high and low intrinsic motivation. *Child Development*, 50, 1207-1210.
- \*Luyten, H., & Lens, W. (1981). The effect of earlier experience and reward contingencies on intrinsic motivation. *Motivation and Emotion*, 5, 25-36.
- Mawhinney, T. C., Dickinson, A. M., & Taylor, L. A. (1989). The use of concurrent schedules to evaluate the effects of extrinsic rewards on "intrinsic motivation." *Journal of Organizational Behavior Management*, 10, 109-129.

- McGraw, K. O. (1978). The detrimental effects of reward on performance: A literature review and a prediction model. In M. R. Lepper & D. Greene (Eds.), *The hidden costs of reward: New perspectives on the psychology of human motivation*. Hillsdale, NJ: Erlbaum.
- \*McGraw, K. O., & McCullers, J. C. (1979). Evidence of a detrimental effect of extrinsic incentives on breaking a mental set. *Journal of Experimental Social Psychology*, 15, 285–294.
- \*McLoyd, V. C. (1979). The effects of extrinsic rewards of differential value on high and low intrinsic interest. *Child Development*, 50, 1010–1019.
- \*Morgan, M. (1981). The overjustification effect: A developmental test of self-perception interpretations. *Journal of Personality and Social Psychology*, 40, 809–821.
- \*Morgan, M. (1983). Decrements in intrinsic interest among rewarded and observer subjects. *Child Development*, 54, 636–644.
- Mossholder, K. W. (1980). Effects of externally mediated goal setting on intrinsic motivation: A laboratory experiment. *Journal of Applied Psychology*, 65, 202–210.
- \*Mynatt, C., Oakley, T., Arkkelin, D., Piccione, A., Margolis, R., & Arkkelin, J. (1978). An examination of overjustification under conditions of extended observation and multiple reinforcement: Overjustification or boredom? *Cognitive Therapy and Research*, 2, 171–177.
- \*Newman, J., & Layton, B. D. (1984). Overjustification: A self-perception perspective. *Personality and Social Psychology Bulletin*, 10, 419–425.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328–346.
- \*Ogilvie, L., & Prior, M. (1982). The overjustification effect in retarded children: Durability and generalizability. *Australia and New Zealand Journal of Developmental Disabilities*, 8, 213–218.
- \*Okano, K. (1981). The effects of extrinsic reward on intrinsic motivation. *Journal of Child Development*, 17, 11–23.
- \*Orlick, T. D., & Mosher, R. (1978). Extrinsic awards and participant motivation in a sport related task. *International Journal of Sport Psychology*, 9, 27–39.
- Overskeid, G., & Svardal, F. (1998). Effect of reward on subjective autonomy and interest when initial interest is low. *Psychological Record*, 46, 319–331.
- \*Pallak, S. R., Costomiris, S., Sroka, S., & Pittman, T. S. (1982). School experience, reward characteristics, and intrinsic motivation. *Child Development*, 53, 1382–1391.
- \*Patrick, C. (1985). *The effect of level of task difficulty on children's intrinsic motivation*. Unpublished doctoral dissertation, Purdue University.
- \*Perry, D. G., Bussey, K., & Redman, J. (1977). Reward-induced decreased play effects: Reattribution of motivation, competing responses, or avoid frustration? *Child Development*, 48, 1369–1374.
- Phillips, J. S., & Freedman, S. M. (1985). Contingent pay and intrinsic task interest: Moderating effects of work values. *Journal of Applied Psychology*, 70, 306–313.
- \*Picek, J. S. (1976). *Effects of reward uncertainty and ability information on attributions of intrinsic motivation*. Unpublished doctoral dissertation, Indiana University.
- \*Pittman, T. S., Cooper, E. E., & Smith, T. W. (1977). Attribution of causality and the overjustification effect. *Personality and Social Psychology Bulletin*, 3, 280–283.
- \*Pittman, T. S., Davey, M. E., Alafat, K. A., Wetherill, K. V., & Kramer, N. A. (1980). Informational versus controlling verbal rewards. *Personality and Social Psychology Bulletin*, 6, 228–233.
- \*Pittman, T. S., Emery, J., & Boggiano, A. K. (1982). Intrinsic and extrinsic motivational orientations: Reward-induced changes in preference for complexity. *Journal of Personality and Social Psychology*, 42, 789–797.
- \*Porac, J. F., & Meindl, J. (1982). Undermining overjustification: Inducing intrinsic and extrinsic task representations. *Organizational Behavior and Human Performance*, 29, 208–226.
- \*Pretty, G. H., & Seligman, C. (1984). Affect and the overjustification effect. *Journal of Personality and Social Psychology*, 46, 1241–1253.
- Pritchard, R. D., Campbell, K. M., & Campbell, D. J. (1977). Effects of extrinsic financial rewards on intrinsic motivation. *Journal of Applied Psychology*, 62, 9–15.
- Quattrone, G. A. (1985). On the congruity between internal states and action. *Psychological Bulletin*, 98, 3–40.
- Reeve, J., & Deci, E. L. (1996). Elements within the competitive situation that affect intrinsic motivation. *Personality and Social Psychology Bulletin*, 22, 24–33.
- \*Reiss, S., & Sushinsky, L. W. (1975). Overjustification, competing responses, and the acquisition of intrinsic interest. *Journal of Personality and Social Psychology*, 31, 1116–1125.
- Reiss, S., & Sushinsky, L. W. (1976). The competing response hypothesis of decreased play effects: A reply to Lepper and Greene. *Journal of Personality and Social Psychology*, 33, 233–245.
- \*Rosenfield, D., Folger, R., & Adelman, H. (1980). When rewards reflect competence: A qualification of the overjustification effect. *Journal of Personality and Social Psychology*, 39, 368–376.
- \*Ross, M. (1975). Salience of reward and intrinsic motivation. *Journal of Personality and Social Psychology*, 32, 245–254.
- \*Ross, M., Karniol, R., & Rothstein, M. (1976). Reward contingency and intrinsic motivation in children: A test of the delay of gratification hypothesis. *Journal of Personality and Social Psychology*, 32, 245–254.
- Rummel, A., & Feinberg, R. (1988). Cognitive evaluation theory: A meta-analytic review of the literature. *Social Behavior and Personality*, 16, 147–164.
- \*Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450–461.
- Ryan, R. M. (1993). Agency and organization: Intrinsic motivation, autonomy and the self in psychological development. In J. Jacobs (Ed.), *Nebraska Symposium on Motivation: Developmental perspectives on motivation* (Vol. 40, pp. 1–56). Lincoln, NE: University of Nebraska Press.
- Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of Personality*, 63, 397–427.
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749–761.
- Ryan, R. M., & Deci, E. L. (1996). When paradigms clash: Comments on Cameron and Pierce's claim that rewards do not undermine intrinsic motivation. *Review of Educational Research*, 66, 33–38.
- Ryan, R. M., & Grolnick, W. S. (1986). Origins and pawns in the classroom: Self-report and projective assessments of individual differences in children's perceptions. *Journal of Personality and Social Psychology*, 50, 550–558.
- Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Ego-involved persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion*, 15, 185–205.
- Ryan, R. M., Kuhl, J., & Deci, E. L. (1997). Nature and autonomy: Organizational view of social and neurobiological aspects of self-regulation in behavior and development. *Development and Psychopathology*, 9, 701–728.
- \*Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, 45, 736–750.
- \*Salancik, G. R. (1975). Interaction effects of performance and money on self-perception of intrinsic motivation. *Organizational Behavior and Human Performance*, 13, 339–351.
- \*Sansone, C. (1986). A question of competence: The effects of competence

- and task feedback on intrinsic interest. *Journal of Personality and Social Psychology*, 51, 918-931.
- \*Sansone, C. (1989). Competence feedback, task feedback, and intrinsic interest: An examination of process and context. *Journal of Experimental Social Psychology*, 25, 343-361.
- \*Sansone, C., Sachau, D. A., & Weir, C. (1989). Effects of instruction on intrinsic interest: The importance of context. *Journal of Personality and Social Psychology*, 57, 819-829.
- \*Sarafino, E. P. (1984). Intrinsic motivation and delay of gratification in preschoolers: The variables of reward salience and length of expected delay. *British Journal of Developmental Psychology*, 2, 149-156.
- Scott, W. E. (1975). The effects of extrinsic rewards on "intrinsic motivation": A critique. *Organizational Behavior and Human Performance*, 14, 117-129.
- Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.
- \*Shanab, M. E., Peterson, D., Dargahi, S., & Deroian, P. (1981). The effects of positive and negative verbal feedback on the intrinsic motivation of male and female subjects. *Journal of Social Psychology*, 115, 195-205.
- Shapira, Z. (1976). Expectancy determinants of intrinsically motivated behavior. *Journal of Personality and Social Psychology*, 34, 1235-1244.
- \*Shiffman-Kaufman, S. E. (1990). *The effects of reward contingency and type of learning experience on intrinsic motivation*. Unpublished doctoral dissertation, City University of New York.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, 52, 270-277.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- \*Smith, A. T. (1980). *Effects of symbolic reward and positive feedback on high and low levels of intrinsic motivation in preschoolers*. Unpublished doctoral dissertation, University of Missouri-Columbia.
- \*Smith, T. W., & Pittman, T. S. (1978). Reward, distraction, and the overjustification effect. *Journal of Personality and Social Psychology*, 36, 565-572.
- \*Smith, W. E. (1975). *The effect of anticipated vs. unanticipated social reward on subsequent intrinsic motivation*. Unpublished doctoral dissertation, Cornell University.
- \*Sorensen, R. L., & Maehr, M. L. (1976). Toward the experimental analysis of "continuing motivation." *Journal of Educational Research*, 69, 319-322.
- \*Staw, B. M., Calder, B. J., Hess, R. K., & Sanderlands, L. E. (1980). Intrinsic motivation and norms about payment. *Journal of Personality*, 48, 1-14.
- \*Swann, W. B., & Pittman, T. S. (1977). Initiating play activity of children: The moderating influence of verbal cues on intrinsic motivation. *Child Development*, 48, 1128-1132.
- Tang, S.-H., & Hall, V. C. (1995). The overjustification effect: A meta-analysis. *Applied Cognitive Psychology*, 9, 365-404.
- \*Taub, S. I., & Dollinger, S. J. (1975). Reward and purpose as incentives for children differing in locus of control expectancies. *Journal of Personality*, 43, 179-195.
- \*Thompson, E. P., Chaiken, S., & Hazlewood, J. D. (1993). Need for cognition and desire for control as moderators of extrinsic reward effects: A Person  $\times$  Situation approach to the study of intrinsic motivation. *Journal of Personality and Social Psychology*, 64, 987-999.
- \*Tripathi, K. N., & Agarwal, A. (1985). Effects of verbal and tangible rewards on intrinsic motivation in males and females. *Psychological Studies*, 30, 77-84.
- \*Tripathi, K. N., & Agarwal, A. (1988). Effect of reward contingency on intrinsic motivation. *Journal of General Psychology*, 115(3), 241-246.
- Utman, C. H. (1997). Performance effects of motivational state: A meta-analysis. *Personality and Social Psychology Review*, 1, 170-182.
- Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94, 3-15.
- \*Vallerand, R. J. (1983). The effect of differential amounts of positive verbal feedback on the intrinsic motivation of male hockey players. *Journal of Sport Psychology*, 5, 100-107.
- Vallerand, R. J., Fortier, M. S., and Guay, F. (1997). Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, 72, 1161-1176.
- \*Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. *Journal of Sport Psychology*, 6, 94-102.
- Vasta, R., Andrews, D. E., McLaughlin, A. M., Stirpe, L. A., & Comfort, C. (1978). Reinforcement effects on intrinsic interest: A classroom analog. *Journal of School Psychology*, 16, 161-168.
- \*Vasta, R., & Stirpe, L. A. (1979). Reinforcement effects on three measures of children's interest in math. *Behavior Modification*, 3, 223-244.
- \*Weinberg, R. S., & Jackson, A. (1979). Competition and extrinsic rewards: Effect on intrinsic motivation and attribution. *Research Quarterly*, 50, 494-502.
- \*Weiner, M. J. (1980). The effect of incentive and control over outcomes upon intrinsic motivation and performance. *Journal of Social Psychology*, 112, 247-254.
- \*Weiner, M. J., & Mander, A. M. (1978). The effects of reward and perception of competency upon intrinsic motivation. *Motivation and Emotion*, 2, 67-73.
- White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, 66, 297-333.
- \*Wicker, F. W., Brown, G., Wiehe, J. A., & Shim, W.-Y. (1990). Moods, goals, and measures of intrinsic motivation. *Journal of Psychology*, 124(1), 75-86.
- Wiersma, U. J. (1992). The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology*, 65, 101-114.
- \*Williams, B. W. (1980). Reinforcement, behavior constraint, and the overjustification effect. *Journal of Personality and Social Psychology*, 39, 599-614.
- Williams, G. C., Grow, V. M., Freedman, Z., Ryan, R. M., & Deci, E. L. (1996). Motivational predictors of weight loss and weight-loss maintenance. *Journal of Personality and Social Psychology*, 70, 115-126.
- \*Wilson, R. L. (1978). *The effect of reward on intrinsic motivation: An integration of dissonance and intrinsic motivation studies*. Unpublished doctoral dissertation, North Carolina State University.
- Wimperis, B. R., & Farr, J. L. (1979). The effects of task content and reward contingency upon task performance and satisfaction. *Journal of Applied Social Psychology*, 9, 229-249.
- \*Yuen, W. C. (1984). *Self-schema, task information, extrinsic reward and intrinsic motivation*. Unpublished doctoral dissertation, Simon Fraser University, Canada.
- \*Zinser, O., Young, J. G., & King, P. E. (1982). The influence of verbal reward on intrinsic motivation in children. *Journal of General Psychology*, 106, 85-91.
- Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). On the importance of self-determination for intrinsically motivated behavior. *Personality and Social Psychology Bulletin*, 4, 443-446.

## Appendix A

## Studies Used in Our Meta-Analyses

A list of each study used in our meta-analyses. A D following the name and date indicates an unpublished dissertation. The second column indicates types of rewards and/or reward contingencies, followed by whether participants were children or undergraduates, followed by whether the dependent measure was free-choice behavior or self-reported interest. (Codes appear in the *Note* to Appendix Table A1.) Finally, we explain

whether our treatment of the study and results differed from Cameron and Pierce's (1994). If a study was coded the same, the same control groups were used in the comparisons, and the effect sizes we reported did not differ from the effect sizes Cameron and Pierce reported by more than 0.10 in either direction, we noted that the study was the same in the two meta-analyses. If there was a difference, we explained what it was.

**Table A1**  
*Studies Used in Our Meta-Analyses Compared With Cameron and Pierce (1994)*

| Study                         | Variables      | Comparison with Cameron & Pierce's (1994) analysis   |
|-------------------------------|----------------|--|
| Amabile et al. (1986), Exp. 1 | E, 1, F, S     | Same. <sup>a</sup>   |
| Amabile et al. (1986), Exp. 3 | E, 2, S        | Same.  |
| R. Anderson et al. (1976)     | V, E, 1, F     | This had multiple no-reward control groups. We selected the one recommended as appropriate by the study's authors and comparable to ones used for other studies in this meta-analysis. C & P <sup>b</sup> used a control group that the authors said was inappropriate, in which the experimenter avoided eye contact with the young children and ignored their attempts to interact, even though there were just the two people in the room. The study's authors said that this condition was uncomfortable, even painful, for both the children and experimenter. Not surprisingly, that group showed free-choice intrinsic motivation that was considerably lower than any other group. |
| S. Anderson & Rodin (1989)    | V, 2, S        | Nearly the same. <sup>c</sup> Both meta-analyses treated the composite dependent variable as self-report.  |
| Arkes (1979)                  | C, 2, F, S     | Same.  |
| Arnold (1976)                 | E, 2, S        | Same.  |
| Arnold (1985)                 | E, C, 2, S     | Same.  |
| Bartelme (1983) D             | P, 2, S        | Excluded, Type I. <sup>d</sup>   |
| Blanck et al. (1984), Exp. 1  | V, 2, F, S     | Same for free-choice; nearly the same for self-report.   |
| Blanck et al. (1984), Exp. 2  | V, 2, F, S     | Excluded, Type II. <sup>e</sup>  |
| Boggiano & Ruble (1979)       | E, P, 1, F     | Excluded, Type II.   |
| Boggiano et al. (1982)        | E, 1, F        | Same.  |
| Boggiano et al. (1985)        | E, C, P, 1, F  | The study's authors crossed reward contingency with salience of reward. They referred to the two reward contingencies as task-contingent and performance-contingent, and C & P coded them that way, treating the task-contingent conditions as engagement-contingent. <sup>f</sup> However, the salience manipulation in the task-contingent condition changed the contingency. In the low-salience group, rewards were given simply for working on the puzzles, which makes the rewards engagement-contingent, but in the high-salience group, rewards were given for each puzzle completed, which makes the rewards completion-contingent.   |
| Brennan & Glover (1980)       | E, 2, F        | This was engagement-contingent because participants got rewards if they "work[ed] with the Soma puzzle for at least 8 minutes" (p. 29), but C & P coded it task-noncontingent. Further, C & P combined two control groups, including one that had not worked on the task for the same amount of time as the rewards group during the experimental period, but we used only the control group that had worked on the task for the same amount of time.  |
| Brewer (1980) D               | E, P, 1, F, S  | Excluded, Type I.  |
| Brockner & Vasta (1981)       | C, 2, F, S     | Same.  |
| Butler (1987)                 | V, 1, S        | Nearly the same.   |
| Calder & Staw (1975)          | C, DT, 2, S    | This study provided monetary rewards for completing a set of puzzles, thus making it completion-contingent, but C & P coded it engagement-contingent. Also, C & P collapsed across interesting and dull tasks. <sup>g</sup>  |
| Chung (1995)                  | E, P, DT, 1, F | Excluded, Type III. <sup>h</sup>   |
| D. S. Cohen (1974) D          | V, P, 2, F, S  | Excluded, Type I.  |
| Crino & White (1982)          | V, 2, F, S     | Same.  |
| Dafoe (1985) D                | N, P, 1, F, S  | Excluded, Type I.  |
| Daniel & Esser (1980)         | P, D, 2, F, S  | In this study, participants were told they could win up to \$2 "depending on how quickly they correctly assembled the 10 puzzles" (p. 568). This conveyed that the rewards depended on doing well relative to a standard and not just on finishing the puzzles. Thus, we coded it performance-contingent, but C & P coded it completion-contingent. Also, C & P collapsed across interesting and dull tasks.   |
| Danner & Lonky (1981), Exp. 2 | V, E, 1, F, S  | Nearly the same.   |
| Deci (1971), Exp. 1           | C, 2, F, S     | Same.  |
| Deci (1971), Exp. 3           | V, 2, F, S     | Same.  |
| Deci (1972a)                  | N, 2, F        | Same.  |

Table A1 (continued)

| Study   | Variables   | Comparison with Cameron & Pierce's (1994) analysis   |
|---|---|--|
| Deci (1972b)<br>Deci et al. (1975)<br>DeLoach et al. (1983)<br>Dimitroff (1984) D<br>Dollinger & Thelen (1978)                    | V, C, 2, F<br>V, 2, F<br>E, 1, F<br>E, 1, F, S<br>V, P, 1, F, S     | Same.<br>Excluded, Type II.<br>Same.<br>Excluded, Type I.<br>This had three tangible-rewards groups, a verbal-rewards group, and a control group. C & P inappropriately collapsed across verbal and tangible rewards, and they did not use the free-choice data.   |
| Earn (1982)   | N, 2, F, S  | Rewards were given "simply for participating in the study" (p. 364), which makes it task-noncontingent, but C & P coded it engagement-contingent.  |
| Efron (1976) D<br>Eisenstein (1985)<br>Enzle et al. (1991)<br>Fabes (1987), Exp. 1  | V, E, P, 2, S<br>U, C, DT, 1, F<br>P, 2, F<br>C, P, 1, F            | Excluded, Type I.<br>Excluded, Type II.<br>Excluded, Type II.<br>Same for the performance-contingent condition. For the other condition, participants were given rewards "when they finished" (p. 8) a block construction, making the condition completion-contingent, but C & P coded it engagement-contingent.   |
| Fabes (1987), Exp. 2  | C, 1, F   | This study used the same procedure as the completion-contingent condition in Fabes (1987, Exp. 1), making it completion-contingent, but C & P coded it engagement-completion.  |
| Fabes et al. (1986)<br>Fabes et al. (1988)  | E, 1, F, S<br>E, 1, F, S  | Excluded, Type II.<br>Same for free-choice, but C & P did not include the self-report. In this study, children selected a face ranging from frown to smile to reflect how much they enjoyed the task, a procedure that is common for obtaining self-report data from young children.   |
| Fabes et al. (1989)<br>Feehan & Enzle (1991), Exp. 2<br>G. S. Goldstein (1980) D  | E, 1, F<br>C, 2, F<br>C, 2, F                                       | Excluded, Type II.<br>Excluded, Type II.<br>Excluded, Type I. This included competition conditions, but we did not use those because competition has a complex effect on intrinsic motivation (Reeve & Deci, 1996).  |
| L. W. Goldstein (1977) D<br>Greene & Lepper (1974)  | V, C, P, 1, F, S<br>U, E, P, 1, F                                   | Excluded, Type I.<br>Same for the two unexpected groups and the engagement-contingent group, but C & P excluded the performance-contingent group.  |
| Griffith (1984) D   | E, DT, 1, F   | Excluded, Type I. To make this study comparable with most others in this meta-analysis, we included only participants who worked in the individual context.  |
| Griffith et al. (1984)  | C, 1, F   | Children were rewarded for finishing reading a passage up to the bookmark, which makes the reward completion-contingent, but C & P coded it engagement-contingent. (The McLoyd [1979] study used the same instructions and C & P did code it completion-contingent.)   |
| Hamner & Foster (1975)  | E, C, DT, 2, S  | Same coding for completion-contingent. In engagement-contingent, participants were paid 75 cents for the 20 minute task, but C & P coded it as task-noncontingent. Also, C & P collapsed across interesting and dull tasks.  |
| Harackiewicz (1979)   | V, E, P, 1, S   | Same for verbal rewards. Nearly the same for engagement-contingent. C & P excluded the two performance-contingent rewards groups.  |
| Harackiewicz & Manderlink (1984)<br>Harackiewicz et al. (1984), Exp. 1<br>Harackiewicz et al. (1984), Exp. 2                      | P, 1, S<br>P, 2, F, S<br>U, P, 2, F, S                              | Same.<br>Same.<br>Same coding, but C & P made an error in the self-report effect size for performance-contingent, showing it as enhancement when in fact it was undermining with a $d = -0.16$ .   |
| Harackiewicz et al. (1984), Exp. 3<br>Harackiewicz et al. (1987)<br>Hitt et al. (1992)<br>Hyman (1985) D<br>Karniol & Ross (1977) | P, 2, F, S<br>P, 1, S<br>E, DT, 2, F, S<br>E, P, 1, F<br>E, P, 1, F | Same.<br>Same.<br>Excluded, Type III.<br>Excluded, Type I.<br>Same except we coded the performance-contingent conditions for whether participants got the maximum rewards with implicit positive feedback or less than maximum rewards with implicit negative feedback.  |
| Kast & Connor (1988)<br>Koestner et al. (1987)<br>Kruglanski et al. (1971)  | V, IC, 1, S<br>V, 2, F, S<br>N, 1, S                                | Excluded, Type II.<br>Same.<br>Rewards were given "because you have volunteered for this study" (p. 610), so they were task-noncontingent, but C & P coded them engagement-contingent.   |
| Kruglanski et al. (1972)<br>Kruglanski et al. (1975), Exp. 1  | U, 1, S<br>C, 1, S  | Same.<br>Participants were rewarded either for the number of coin flips they guessed correctly or for the number of block constructions they completed correctly, making the study completion-contingent, but C & P coded it performance-contingent. It explored moderation by endogenous versus exogenous rewards.  |
| Kruglanski et al. (1975), Exp. 2  | P, 1, S   | There were two reward groups and two control groups. In one pair, people worked on a stock market game and earned cash after each trial for good investments. The control group was the same as the experimental group except they were told they had to give back their earnings, so it was not a reasonable no-reward control group. In the other pair of conditions, money was not mentioned to the no-reward control group. We excluded the pair of conditions without a proper control group, but C & P collapsed across the two pairs of conditions. |
| Lee (1982) D<br>Lepper et al. (1973)  | P, 2, F, S<br>U, E, 1, F  | Excluded, Type I.<br>Same coding. Same effect sizes for engagement-contingent. C & P made an error in calculating the effect size for unexpected rewards.  |
| Lepper et al. (1982), Exp. 3  | E, 1, F   | Excluded, Type II.   |

(table continues)

Table A1 (*continued*)

| Study  | Variables   | Comparison with Cameron & Pierce's (1994) analysis   |
|--|---|--|
| Liberty (1986), Exp. 1 D<br>Liberty (1986), Exp. 2 D<br>Loveland & Olley (1979)<br>Luyten & Lens (1981)  | C, 2, F, S<br>C, 2, F, S<br>E, DT, 1, F<br>C, P, 2, F, S  | Excluded, Type I.<br>Excluded, Type I.<br>Same coding, but C & P collapsed across interesting and dull tasks.<br>Same for performance-contingent. In the other rewards condition, participants were paid after each of three puzzles they solved, so it was completion-contingent, but C & P coded it as engagement-contingent.  |
| McGraw & McCullers (1979)<br>McLoyd (1979)<br>Morgan (1981), Exp. 1<br>Morgan (1981), Exp. 2<br>Morgan (1983), Exp. 1<br>Morgan (1983), Exp. 2<br>Mynatt et al. (1978)<br>Newman & Layton (1984)<br>Ogilvie & Prior (1982)<br>Okano (1981), Exp. 1<br>Okano (1981), Exp. 2<br>Orlick & Mosher (1978) | C, 2, S<br>C, DT, 1, F<br>E, 1, F, S<br>E, 1, F, S<br>E, 1, F, S<br>E, 1, F, S<br>E, DT, 1, F<br>E, DT, 1, F<br>E, 1, F<br>E, 1, F, S<br>N, E, 1, F, S<br>V, U, P, 1, F | Same.<br>Coded the same, but C & P collapsed across interesting and dull tasks.<br>Same on free-choice; nearly the same on self-report.<br>Same.<br>Same on free-choice; nearly the same on self-report.<br>Same.<br>Coded the same, but C & P collapsed across interesting and dull tasks.<br>Excluded, Type II.<br>Same.<br>Excluded, Type II.<br>Excluded, Type II.<br>Same coding for verbal and unexpected. In performance-contingent, children got rewards "if you do a good job today and tomorrow on the balance board" (p. 31), but C & P coded it as completion-contingent. There were discrepancies in the effect sizes.  |
| Pallak et al. (1982)<br>Patrick (1985) D<br>Perry et al. (1977)<br>Picek (1976) D<br>Pittman et al. (1977)   | V, U, P, 1, F<br>E, P, 1, F, S<br>E, 1, F, S<br>E, P, 2, F, S<br>P, 2, F, S   | Same for verbal and unexpected. C & P did not report how they coded the tangible expected rewards condition, which was performance-contingent.<br>Excluded, Type I.<br>Excluded, Type II.<br>Excluded, Type I.<br>Same coding, but C & P used only self-report. We also used free-choice persistence, calculated as the number of trials.  |
| Pittman et al. (1980)<br>Pittman et al. (1982), Exp. 1<br>Pittman et al. (1982), Exp. 2<br>Porac & Meindl (1982)   | V, IC, 2, F<br>N, E, 1, F<br>E, 1, F<br>C, 2, F   | Same except that C & P did not do an analysis of informational versus controlling positive feedback.<br>Same codings and nearly the same free-choice effects. C & P imputed a self-report value of 0.00, but participants were not asked how interesting or enjoyable they found the activity.<br>Nearly the same.<br>C & P coded this engagement-contingent, but participants received \$1.50 for each puzzle solved. C & P reported a comparison for 40 experimental and 20 control participants, but there were only 50 participants in the study. We calculated the reward effect size based on a comparison of the rewarded groups with neutral and extrinsic mind-sets versus the nonrewarded groups with neutral and extrinsic mind-sets, because that comparison provided corresponding reward versus no-reward conditions.                      |
| Pretty & Seligman (1984), Exp. 1<br>Pretty & Seligman (1984), Exp. 2<br>Reiss & Sushinsky (1975), Exp. 1<br>Rosenfield et al. (1980)   | V, U, E, 2, F, S<br>U, E, 2, F, S<br>E, 1, F<br>P, 2, F, S  | Same for unexpected and engagement-contingent. Nearly the same for verbal on free-choice.<br>Same.<br>Same.<br>This study had performance-contingent, completion-contingent, and task-noncontingent groups, and a control group with feedback comparable to that in performance-contingent. There was no appropriate control group for completion-contingent or task-noncontingent. The study also crossed tangible rewards with positive versus negative feedback. C & P reported a verbal effect for positive versus negative feedback, and then they collapsed across feedback to examine tangible-reward effects. We did a moderator analysis of rewards signifying positive versus negative feedback. C & P listed a performance-contingent self-report $d = 2.80$ , but the correct $d$ was 0.22. For free-choice, there was a modest discrepancy. |
| Ross (1975), Exp. 1<br>Ross (1975), Exp. 2<br>Ross et al. (1976)   | E, 1, F, S<br>E, 1, F, S<br>N, E, 1, F  | Same for free-choice; C & P did not include self-report.<br>Nearly the same for free-choice; C & P did not include self-report.<br>Same for engagement-contingent. In the other group, children were rewarded for waiting, which is task-noncontingent, but C & P coded it engagement-contingent.  |
| Ryan (1982)<br>Ryan et al. (1983)  | IC, 2, F<br>V, E, P, IC, 2, F, S  | We included this study only in the supplemental meta-analysis of informational versus controlling verbal rewards. C & P excluded it.<br>Same on verbal and engagement-contingent. There were two performance-contingent groups, one informational and one controlling. There were three no-reward control groups, one with informational positive feedback, one with controlling positive feedback, and one with no feedback. We compared performance-contingent with both comparable-feedback controls and no-feedback controls in the moderator analyses. C & P did only the comparable-feedback comparisons. Also, C & P did not do an informational-controlling comparison.  |
| Salancik (1975)  | P, 2, F, S  | Same coding. C & P collapsed across positive and negative feedback conditions, but we did a moderator analysis for positive versus negative.   |
| Sansone (1986)<br>Sansone (1989)<br>Sansone et al. (1989)  | V, 2, S<br>V, 2, S<br>V, 2, S   | Same.<br>Same.<br>Same.  |

(table continues)

Table A1 (continued)

| Study                          | Variables        | Comparison with Cameron & Pierce's (1994) analysis  |
|--------------------------------|------------------|---|
| Sarafino (1984)                | E, 1, F, S       | Same.   |
| Shanab et al. (1981)           | V, 2, F, S       | Same.   |
| Shiffman-Kaufman (1990) D      | E, P, 1, F, S    | Excluded, Type I. For comparability with other studies, we used only data from the 10-day assessments.  |
| A. T. Smith (1980) D           | E, DT, 1, F      | Excluded, Type I. In this study, there was also a condition called positive feedback, but the statements were not competence feedback.  |
| T. W. Smith & Pittman (1978)   | P, 2, F, S       | Same for self-report. C & P imputed a score of 0.00 for free-choice performance, even though means and significance tests were reported.  |
| W. E. Smith (1975) D           | V, U, P, 2, F, S | Excluded, Type I.   |
| Sorensen & Maehr (1976)        | C, 1, F          | Excluded, Type II.  |
| Staw et al. (1980)             | C, 2, S          | Participants got a \$1 reward for completing 15 puzzles, making the study completion-contingent, but C & P coded it engagement-contingent.  |
| Swann & Pittman (1977), Exp. 1 | N, E, 1, F       | Same.   |
| Swann & Pittman (1977), Exp. 2 | E, 1, F          | There were two engagement-contingent groups, an engagement-contingent plus verbal-rewards group, and two no-reward control groups. There was not a control group for the engagement plus verbal group. We compared the two engagement with the two control groups, but C & P used all three reward groups.  |
| Taub & Dollinger (1975)        | P, 2, S          | Same.   |
| Thompson et al. (1993)         | E, 2, F          | Excluded, Type III.   |
| Tripathi & Agarwal (1985)      | V, E, 2, F, S    | Nearly the same.  |
| Tripathi & Agarwal (1988)      | E, P, 2, F, S    | Same for engagement-contingent on free-choice. For performance-contingent, there were two tasks, with free-choice data reported for only one. Both we and C & P used the data for the one task and assigned $d = 0.00$ for the other, but C & P averaged the effects, whereas we combined them meta-analytically. In the self-report data, C & P combined the engagement and performance conditions, so it is unclear which analysis they were used in. |
| Vallerand (1983)               | V, 1, S          | Same.   |
| Vallerand & Reid (1984)        | V, 2, S          | Same.   |
| Vasta & Stirpe (1979)          | C, 1, F          | This study had pre-post data for a rewards group and a control group. C & P did pre-post analyses for the rewards group and ignored the control group. We compared the rewards group with the control group with pre-post analyses. We coded it completion-contingent, but C & P did not code it.   |
| Weinberg & Jackson (1979)      | P, 2, S          | Same.   |
| Weiner (1980)                  | C, 2, F, S       | Participants received \$.25 for each anagram completed, which makes the study completion-contingent, but C & P coded it performance-contingent.   |
| Weiner & Mander (1978)         | E, P, 2, F, S    | Same.   |
| B. W. Williams (1980)          | E, 1, F, S       | Same.   |
| Wilson (1978) D                | E, DT, 2, F, S   | Excluded, Type I.   |
| Wimperis & Farr (1979)         | N, C, 2, S       | In one group, participants received \$1.75 for being in the study, making it task-noncontingent, but C & P coded it engagement-contingent. In the other, participants were paid for each model or subunit completed, making it completion-contingent, but C & P coded it performance-contingent.  |
| Yuen (1984) D                  | E, 2, F, S       | Excluded, Type I.   |
| Zinser et al. (1982)           | V, 1, F          | Same.   |

Note. D = unpublished dissertation; V = verbal rewards; U = unexpected tangible rewards; N = task-noncontingent rewards; E = engagement-contingent rewards; C = completion-contingent rewards; P = performance-contingent rewards; DT = dull-task condition included in study and used in supplemental meta-analysis; IC = informational versus controlling comparison was made in supplemental meta-analysis; 1 = participants were children; 2 = participants were undergraduates; F = free-choice dependent measure was used; S = self-report measure was used.

<sup>a</sup> Same means that Cameron and Pierce (1994) and we coded the study the same, used the same control groups, and found effects sizes that did not differ from each other by more than 0.10 in either direction. <sup>b</sup> C & P refers to Cameron and Pierce (1994). <sup>c</sup> Nearly the same means the studies were coded the same, and the same control groups were used, but the effect sizes differed by more than 0.10, probably due to differences in estimation of standard deviations. If the discrepancy is large, we make note of that. <sup>d</sup> Excluded, Type I refers to dissertations, and Cameron and Pierce (1994) excluded all dissertations. <sup>e</sup> Excluded, Type II refers to studies that Cameron and Pierce (1994) excluded for no apparent reason. <sup>f</sup> Cameron and Pierce (1994) did not use the term engagement-contingent. When we say they coded a reward engagement-contingent, it means that they coded it as both "task-contingent" and what they referred to as not contingent using a behavioral definition. Because the intersection of those two codes is equivalent to our engagement-contingent code, we say that they coded it as engagement-contingent to minimize confusion for the reader. Similarly, they did not use the term completion-contingent, but what they coded as both "task-contingent" and contingent using a behavioral definition is equivalent to what we call completion-contingent. <sup>g</sup> These studies used both interesting and uninteresting tasks. We excluded the uninteresting tasks from the primary meta-analyses and included them in the supplemental meta-analysis concerned with initial task interest. Cameron and Pierce (1994) collapsed across the interesting and dull tasks even though it has been firmly established in the literature that initial task interest interacts with reward effects. <sup>h</sup> Excluded, Type III refers to studies that Cameron and Pierce (1994) excluded because they were published after Cameron and Pierce's cutoff date.

(Appendices continue)

## Appendix B

### Studies Excluded From Our Meta-Analyses

Studies included in the meta-analyses by Eisenberger and Cameron (1996) and Cameron and Pierce (1994) that we excluded, with the explanations for why we excluded them.

**Table B1**  
*Studies We Excluded*

| Study                      | Reason for exclusion  |
|----------------------------|---|
| Boal & Cummings (1981)     | We excluded this because all participants received monetary payments (hourly wages) so there was not a no-rewards control group. For more detail, see Footnotes 1 and 6 to main text.   |
| Boggiano & Hertel (1983)   | We excluded this because participants did not work on the target task before the dependent variable was assessed. Cameron and Pierce (1994) coded it unexpected and reported a $d = 0.02$ , which is as we would expect.  |
| Freedman & Phillips (1985) | We excluded this because the interest level of the task was below the midpoint of the scale, so it did not examine reward effects on interesting tasks.   |
| Hom (1987)                 | Very little detail was presented about methods or results. Cameron & Pierce (1994) indicated that they could not tell if the reward was expected or unexpected or what type of reward contingency was used, so it is unclear whether they actually included the study in any analyses.  |
| Shapira (1976)             | A study of preferences for level of challenge under reward versus no reward. We excluded it because rewarded participants worked on much easier puzzles than did nonrewarded participants.  |
| Wicker et al. (1990)       | This study was not designed to evaluate reward effects on intrinsic motivation. It concerned the correlations between free-choice persistence and various mood variables, so the rewards manipulation was incidental. No cell means were reported for any variables, and there was no mention of whether any between-group tests were done or whether rewards affected any variables. |

*Note.* Cameron and Pierce (1994) did a separate, supplemental meta-analysis of studies with single-subject designs that did not have no-reward control groups. We did not include those studies because they did not meet the inclusion criterion of having a no-reward control group. The studies were Davidson and Bucher (1978); Feingold and Mahoney (1975); Mawhinney, Dickinson, and Taylor (1989); and Vasta, Andrews, McLaughlin, Stirpe, and Comfort (1978).

Received December 11, 1997

Revision received February 16, 1999

Accepted February 17, 1999 ■