The Behavior Analyst 2001, 24, 1–44 No. 1 (Spring)

Pervasive Negative Effects of Rewards on Intrinsic Motivation: The Myth Continues

Judy Cameron, Katherine M. Banko, and W. David Pierce
University of Alberta

A major concern in psychology and education is that rewards decrease intrinsic motivation to perform activities. Over the past 30 years, more than 100 experimental studies have been conducted on this topic. In 1994, Cameron and Pierce conducted a meta-analysis of this literature and concluded that negative effects of reward were limited and could be easily prevented in applied settings. A more recent meta-analysis of the literature by Deci, Koestner, and Ryan (1999) shows pervasive negative effects of reward. The purpose of the present article is to resolve differences in previous meta-analytic findings and to provide a meta-analysis of rewards and intrinsic motivation that permits tests of competing theoretical explanations. Our results suggest that in general, rewards are not harmful to motivation to perform a task. Rewards given for low-interest tasks enhance free-choice intrinsic motivation. On high-interest tasks, verbal rewards produce positive effects on free-choice motivation and self-reported task interest. Negative effects are found on high-interest tasks when the rewards are tangible, expected (offered beforehand), and loosely tied to level of performance. When rewards are linked to level of performance, measures of intrinsic motivation increase or do not differ from a nonrewarded control group. Overall, the pattern of results indicates that reward contingencies do not have pervasive negative effects on intrinsic motivation. Theoretical and practical implications of the findings are addressed.

Key words: meta-analysis, rewards, reinforcement, intrinsic motivation, intrinsic interest

Most parents, educators, and behavior analysts would agree that the ideal student is one who performs academic tasks at a high level, shows high interest and involvement in school activities, is willing to take on challenging assignments, and is a self-motivated learner. To instill interest and to heighten student performance, many practitioners implement reward and incentive systems in educational settings. In recent years, the wisdom of this practice has been debated in literature reviews, textbooks, and the popular media. Many writers and researchers claim that giving students high grades, prizes, money, and even praise for engaging in an activity may be effective in getting students to perform a task, but performance and interest are maintained only as long as the rewards keep coming. In other words, rewards are said to undermine intrinsic motivation. This premise is based on the view that when individuals like what they are doing, they experience feelings of competence and self-determination. When students are given a reward for performance, the claim is that they begin to do the activity for the external reward rather than for intrinsic reasons. As a result, perceptions of competence and self-determination are said to decrease and motivation to perform the activity declines.

Those who decry the use of rewards support their position by citing experimental studies on rewards and intrinsic motivation conducted in social psychology and education. Since the 1970s, dozens of experiments have been designed to assess the impact of rewards on intrinsic motivation. A cursory examination of the studies, however, reveals a mixed set of findings. That is, in some studies, extrinsic re-

This work was supported by a research grant from the Social Sciences and Humanities Research Council of Canada.

Correspondence concerning this article should be addressed to Judy Cameron, Department of Educational Psychology, 6-102 Education North, University of Alberta, Edmonton, Alberta, T6G 2G5 Canada.

wards produce negative effects on measures of intrinsic motivation. Other studies find positive effects of reward; still others show no effect. A number of reviewers have noted the contradictory nature of the findings and have attempted to identify the conditions under which extrinsic rewards produce decrements on measures of intrinsic motivation (Bates, 1979; Bernstein, 1990; Carton, 1996; Dickinson, 1989; Flora, 1990; Morgan, 1984).

In 1994, Cameron and Pierce published a meta-analysis of 96 experimental studies on the topic (with additional analyses by Eisenberger & Cameron, 1996). Based on their results, they argued that negative effects of reward were minimal and could be easily prevented in applied settings. The research and recommendations made by Cameron and Pierce and by Eisenberger and Cameron generated considerable debate (Hennessey & Amabile, 1998; Kohn, 1996; Lepper, 1998; Lepper, Keavney, & Drake, 1996; Ryan & Deci, 1996; Sansone & Harackiewicz, 1998) and seemingly spurred the publication of a new meta-analysis on the topic. Deci, Koestner, and Ryan (1999) presented a meta-analysis that claimed to support the view that rewards have pervasive negative effects on intrinsic motivation.

Deci et al. (1999) identified 128 experiments on rewards and intrinsic motivation, including 20 unpublished studies from doctoral dissertations. They outlined a number of concerns they had with the meta-analyses conducted by Cameron and Pierce (1994) and Eisenberger and Cameron (1996). Deci et al.'s meta-analysis was designed to rectify these concerns, to test cognitive evaluation theory, and to provide a more comprehensive review of the literature. Their findings supported cognitive evaluation theory and, in general, rewards were found to have a substantial negative effect on intrinsic motivation. Deci et al. concluded that ‘‘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’’ (p. 659).

The assertion that rewards decrease intrinsic motivation has captured the attention of cognitive researchers, practitioners, and the general public because such a claim (a) seems to offer an empirical basis for psychological theories that assume that self-determination and freedom from control are fundamental human motives, (b) appears to question basic behavioral conceptions of human nature, and (c) suggests that rewards used in schools, hospitals, the workplace, and so on are more harmful than beneficial. A reviewer of this manuscript suggested that the claim that rewards are harmful may be attractive to some practitioners and educators because detecting and rewarding performance improvements is hard work and the negative effect claim relieves us of a difficult and demanding task.

Clearly, Deci et al.'s (1999) finding of general negative effects of reward has important theoretical and practical implications and calls for a careful analysis of contradictory empirical claims. In this article, we argue that pervasive negative effects of reward are not a necessary outcome of a meta-analysis of this literature. We contend that a careful examination of Deci et al.'s meta-analysis reveals several conceptual and methodological shortcomings. The disparate conclusions of the two major meta-analyses on the effects of reward on intrinsic motivation (Cameron & Pierce, 1994; Deci et al.) suggest the value of correcting the flaws in each and building on their strengths to draw more definitive conclusions. In this article, we offer a reanalysis of the effects of rewards on intrinsic motivation. Our reanalysis is informed by a consideration of Deci et al.'s decisions and procedures. In addition, the concerns raised by Deci et al. about our previous research are addressed. The purpose of the present article is to resolve differences in previous meta-analytic findings and to provide a meta-analysis of rewards and intrinsic motivation that permits tests of competing theoretical explanations.

We begin with a general description of the experiments conducted on rewards and intrinsic motivation. This is followed by a brief description of the procedure and logic of meta-analysis. The meta-analyses by Cameron and Pierce (1994), Eisenberger and Cameron (1996), and Deci et al. (1999) are described, and criticisms of each are presented. We then provide a detailed account of how our reanalysis is designed to resolve differences between Deci et al.'s and our earlier reviews of this literature. Results of our new meta-analysis are presented, and differences between our findings and previous reviews are explained. Finally, our discussion focuses on theoretical and practical implications of the findings.

THIRTY YEARS OF RESEARCH ON REWARDS AND INTRINSIC MOTIVATION

The term *intrinsic motivation* is generally understood in contrast to *extrinsic motivation.* Intrinsically motivated behaviors are those in which there is no apparent reward except with the activity itself (Deci, 1975). Extrinsic motivation, on the other hand, is said to occur when an activity is rewarded by incentives not inherent in the task. Although these terms have been criticized and debated (e.g., Bandura, 1986; Dickinson, 1989; Flora, 1990), they are accepted by many researchers. The distinction between intrinsic and extrinsic motivation led psychologists to speculate about the relation between these two sources. One view was that intrinsic and extrinsic motivation combined in an additive fashion to produce overall motivation. For example, in work settings, organizational psychologists argued that optimal performance would occur when jobs were interesting and challenging and employees were externally rewarded (e.g., with money) for their work (Porter & Lawler, 1968; Vroom, 1964). Other theorists challenged the additive assumption, suggesting instead that extrinsic rewards might interfere with intrinsic motivation (DeCharms, 1968).

The idea that extrinsic rewards could disrupt intrinsic motivation instigated a series of experiments carried out in the early 1970s (Deci, 1971; Lepper, Greene, & Nisbett, 1973). In the initial studies, researchers tested the hypothesis that external rewards would undermine intrinsic motivation either by subverting feelings of competence and self-determination or by deflecting the source of motivation from internal to external causes. Intrinsic motivation was inferred from changes in time spent on an activity once rewards were removed, performance during the nonrewarded phase, or expressed task interest. When rewards were found to lower time on task, performance, or interest, the researchers claimed that rewards undermined intrinsic motivation. Results from the early studies appeared to offer some support for the undermining hypothesis. That is, when individuals were promised a material reward, their performance, time on task, and interest decreased once the reward was no longer forthcoming. Because of the implications for education, business, and the psychology of motivation, the early findings led to a great deal of research on the topic.

Since the 1970s, over 100 experiments have been performed to investigate alleged undermining effects of rewards. The vast majority of the studies on rewards and intrinsic motivation have been conducted using a between-groups design. In a typical study, participants are presented with an interesting task (e.g., solving and assembling puzzles, drawing with magic markers, playing word games). Participants are rewarded with money or grades, candy, praise, good-player certificates, and so forth for performing the activity. Rewards are tangible (e.g., money, candy, gold stars) or verbal (e.g., praise, approval, positive feedback). In addition, the rewards may be offered beforehand (expected reward) or presented unexpectedly after the activity (unexpected reward). In some experiments, reward is offered simply for doing an activity; in other studies the rewards are given for completing a task or for each puzzle or unit solved. In a number of experiments, the rewards are offered for meeting or exceeding a specific standard. Participants in a control condition engage in the activity without receiving a reward.

The reward intervention is usually conducted over a 10-min to 1-hr period. Rewarded and nonrewarded groups are then observed during a nonreward period (typically, 2 min to 1 hr) in which participants are free to continue performing the target task or to engage in some alternative activity. The time participants spend on the target activity during this nonreward phase, their performance on the task during the free-choice period, or self-reported task interest are used as measures of intrinsic motivation. If rewarded participants spend less free time on the activity, perform at a lower level, or express less task interest than nonrewarded participants, reward is said to undermine intrinsic motivation.

The findings from the studies on rewards and intrinsic motivation have been diverse (positive, negative, and no effects have been reported). Nonetheless, the results from these studies are often cited as evidence that rewards and positive reinforcement can backfire (e.g., Kohn, 1993). External rewards are said to be controlling and to interfere with a basic human desire for self-determination.

Because the detrimental effects of rewards have been interpreted as a challenge to behavioral conceptions of human nature and to the benefits of behavioral technology for education and business, a few behaviorally oriented researchers have used single-subject designs to assess the generality of the findings. In this type of study, participants serve as their own controls. Measures such as time on task are taken over a number of sessions in a baseline phase, reinforcement procedures are implemented over several sessions, and finally, reward is withdrawn and time on task is assessed on repeated occasions. An increase or decrease in intrinsic motivation is measured by the difference in time spent on the task between baseline and postreinforcement phases. In the five studies employing this type of design (Davidson & Bucher, 1978; Feingold & Mahoney, 1975; Mawhinney, Dickinson, & Taylor, 1989; Skaggs, Dickinson, & O'Connor, 1992; Vasta, Andrews, McLaughlin, Stirpe, & Comfort, 1978), participants' performance during the postreward phase either exceeded or remained at the same level as performance in the prereward sessions. In other words, when the rewards were shown to function as reinforcement and multiple-trials procedures were used, there was no evidence of a decremental effect of reward.

Those who argue that rewards decrease intrinsic motivation are critical of the single-subject designs. For example, Deci et al. (1999) state that one can conclude very little from the single-subject designs because there are too few participants and none of the studies had control groups. Instead, claims about negative effects of reward are based on results from the group-design studies. As noted, however, the findings from such studies have not been entirely clear-cut. Although most researchers have found that verbal rewards do not decrease measures of intrinsic motivation, the results with tangible rewards have been more contradictory. To understand such diverse effects, Cameron and Pierce (1994), Eisenberger and Cameron (1996), and Deci et al. used the methodology of meta-analysis to assess the group-design experiments and to determine when and under what conditions rewards have detrimental effects on task performance and interest. Despite the seeming objectivity of this technique, these meta-analytic reviews reached markedly different conclusions. Cameron and Pierce and Eisenberger and Cameron reported minimal negative effects of tangible reward, whereas Deci et al. found tangible rewards to be detrimental under a wide range of conditions.

Although the usefulness of meta-analysis and statistical testing in general has been questioned by behavioral researchers (e.g., see Baron & Derenne, 2000; Derenne & Baron, 1999), research summaries based on meta-analyses have become valued sources of information for both policy makers and researchers. Deci et al.'s (1999) meta-analytic finding of general negative effects of reward has important implications. Thus, to understand why the meta-analyses by Cameron and Pierce (1994) and Deci et al. resulted in different findings, it is important to be familiar with the technique and logic of meta-analysis. The meta-analytic procedures described below are based on Hedges and Olkin (1985); these were the basic procedures used by Cameron and Pierce and by Deci et al.

THE TECHNIQUE AND LOGIC OF META-ANALYSIS

Meta-analysis is a technique for combining the results of a large number of studies on the same topic. It involves combining data from conceptually related studies to reach generalizations based on statistical criteria. Quantitative analyses, similar to meta-analysis, have been conducted on single-subject designs (e.g., see Kollins, Newland, & Critchfield, 1997); however, meta-analysis is typically used with between-groups designs in which a treatment group (e.g., a rewarded group) is compared to a control group (nonrewarded group) on a common dependent measure (intrinsic motivation). The goals of a meta-analysis are to establish the relation between independent and dependent variables (in this case, the relation between rewards and intrinsic motivation) and to determine what factors moderate or alter the magnitude of the relation (e.g., type of reward, reward contingency). Conducting a meta-analysis entails specifying the criteria for including and excluding studies, collecting all experiments that meet the criteria, and coding the studies.

Once all relevant studies are identified, the statistical results of each study are transformed into a measure called an effect size. An effect size is found by converting the findings from each study into a standard deviation unit. In the rewards and intrinsic motivation literature, an effect size indicates the extent to which the experimental group (rewarded group) and the control group (nonrewarded group) differ in the means on measures of intrinsic motivation (e.g., free time on task after rewards are removed, task interest). In its simplest form, the effect size (g) is the difference between the means of the rewarded group and the nonrewarded control group divided by the pooled standard deviation of this difference. In a meta-analysis, the effect size from each study, rather than the individual participants within a study, becomes the unit of analysis. If the effect sizes from all the studies present a random pattern, they will hover around zero, indicating no evidence for an effect. On the other hand, the effect sizes may cluster in a positive or negative direction, indicating that something is going on.

One problem in meta-analysis arises when studies do not provide enough information to calculate effect sizes. When means and standard deviations are not available, effect sizes can be calculated from t tests, F statistics, and p values (see Hedges & Becker, 1986). However, in some cases, there may still be insufficient information to obtain an effect size. The meta-analyst can contact the researchers and try to obtain the missing data. When the data cannot be procured, the study can be excluded from the analyses or assigned an effect size of 0.00 (indicating no difference between experimental and control groups). It has been argued that including zero effect sizes is a conservative strategy; if a significant effect is detected in spite of the inclusion of zeros, the contention is that the results would not be altered if missing data were available (for a discussion of this issue, see Light & Pillemer, 1984). On the other hand, if one's bias is toward no effect (i.e., we are satisfied if the treatment is not harmful), including zeros favors this conclusion. One strategy for dealing with this issue is to conduct the analyses with zeros included and excluded.

After effect sizes (g) are calculated for each relevant study, an overall mean effect size ($d+$) is obtained. First, g is converted to d by correcting for bias (g is an overestimation of the population effect size, particularly for small samples; see Hedges, 1981). The overall mean effect size is obtained by weighting each effect size by the reciprocal of its variance and averaging the weighted d. This procedure gives more weight to effect sizes that are more reliably estimated. The calculation of mean effect sizes provides a significance test (whether the value differs significantly from 0.00) and a 95% confidence interval (CI) (when the CI contains 0.00, the results suggest that there is no evidence of a statistically significant effect).

In a hierarchical meta-analysis, all studies are included in an overall analysis. The researcher then searches for moderator variables. The studies are broken out by one key moderator, then another, and so on. The moderators that the researcher chooses to examine may be based on theoretical considerations or on differences between the studies (e.g., different procedures used in the studies, different characteristics of the samples used, year of publication, etc.).

Hedges and Olkin (1985) recommend using homogeneity tests to ascertain whether a moderator analysis is necessary. Essentially, the procedure is to use a chi-square statistic, Q, with $K - 1$ degrees of freedom, where K is the number of effect sizes. The null hypothesis is that the effect sizes are homogeneous (i.e., effect sizes in a given analysis are viewed as values sampled from a single population; variation in effect sizes among studies is merely due to sampling variation). When Q is statistically significant, the implication is that moderator analyses should be conducted. The original set of studies is then broken into subsets until the chi-square statistics within the subgroups are nonsignificant. When the researcher has exhausted potential moderators and homogeneity is still not obtained, outliers (studies with extreme effect-size values) are examined independently and the analysis is conducted with outliers removed.

DIFFERENCES BETWEEN CAMERON AND PIERCE'S (1994) AND DECI ET AL.'S (1999) META-ANALYSES

Although Deci et al. (1999) and Cameron and Pierce (1994) used the same meta-analytic procedures to evaluate the research on rewards and intrinsic motivation, their results differed. Cameron and Pierce conducted a hierarchical meta-analysis of the rewards and intrinsic motivation literature. Studies were included if they had a rewarded group and a nonrewarded control group and if they used one of the two main measures of intrinsic motivation (free time on the task after the reward was removed or self-reported task interest). The effects of reward on the two dependent measures (free time and task interest) were assessed separately. When a study did not provide enough information to calculate an effect size, it was not included in the analyses.

Cameron and Pierce (1994) were first interested in whether rewards, overall, produced negative effects on measures of intrinsic motivation. Their findings indicated no overall negative effects on either measure of intrinsic motivation. However, the set of effect sizes was significantly heterogeneous; thus, the researchers conducted a number of moderator analyses to determine when and under what conditions rewards produced negative effects. Rewards were broken down by reward type (tangible and verbal). Tangible rewards were subdivided into expected and unexpected, and expected tangible rewards were further separated by the reward contingency. Cameron and Pierce used a behavioral framework to categorize rewards by reward contingency; in addition, they used the categories suggested by Deci and Ryan's (1985) cognitive evaluation theory framework. Their results indicated negative effects on the free-time measure only when the rewards were tangible, expected, and not contingent on meeting a performance standard. The same findings were reported by Eisenberger and Cameron (1996), who carried out some additional analyses of reward contingencies.

Deci et al. (1999) suggested that Cameron and Pierce's (1994) and Eisenberger and Cameron's (1996) failure to detect more pervasive negative effects was due to methodological inadequacies. Specifically, they criticized Cameron and Pierce and Eisenberger and Cameron for the following: (a) collapsing across tasks with high and low initial interest and omitting a moderator analysis of initial task interest, (b) including a study that used an inappropriate control group (Boal & Cummings, 1981), (c) omitting studies or data as outliers rather than attempting to isolate moderators, (d) omitting studies that were published during the period covered by their meta-analysis, (e) omitting unpublished doctoral dissertations, and (f) misclassifying studies into reward contingencies as defined by cognitive evaluation theory.

To rectify these issues in their recent meta-analysis, Deci et al. (1999) excluded the study by Boal and Cummings (1981), included studies that were missed in the previous meta-analyses, and included unpublished doctoral dissertations. In addition, in contrast to Cameron and Pierce (1994), for studies with insufficient information to calculate effect sizes, Deci et al. imputed effect sizes of 0.00 and included these in each of their analyses.

In terms of initial task interest, Deci et al. (1999) noted that "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" (p. 633). Given that cognitive evaluation theory has little to say about the effects of rewards on low-interest tasks, Deci et al.'s meta-analysis focused on reward effects on high-interest tasks. Studies or conditions within studies were included only if the tasks used were measured or defined to be initially interesting; studies or conditions within studies were excluded if the tasks used were measured or defined as initially uninteresting.

Thus, Deci et al.'s (1999) meta-analysis began with the overall effects of rewards on intrinsic motivation for tasks of initial high interest only. Deci et al. analyzed the effects of reward on measures of self-reported task interest and free-choice intrinsic motivation. Their free-choice measure included time spent on a task after rewards were removed. When a time measure was not reported in a study, Deci et al. used measures of task persistence during the free-choice period (e.g., number of trials initiated in a labyrinth game, number of balls played in a pinball game, number of successes on a task). Hence, Deci et al.'s analysis of the free-choice measure was broader than the analysis by Cameron and Pierce (1994), who used only studies that assessed time measures.

On tasks of high initial interest, Deci et al. (1999) found a significant negative effect of rewards on the free-choice measure and a non-significant effect on the self-report measure. Both mean effect sizes were heterogeneous. To obtain homogeneity at each level of analysis, Deci et al. tested a number of moderator variables. When homogeneity could not be obtained, Deci et al. followed the procedure used by Cameron and Pierce (1994) and identified and removed outliers. First, Deci et al. tested whether verbal versus tangible rewards were a moderator. Verbal rewards were found to increase free-choice intrinsic motivation for college students (a nonsignificant effect was found for children) and to enhance task interest for both children and college students. Tangible rewards produced negative effects on both the free-choice and self-report measures. In accord with Cameron and Pierce, tangible rewards were broken down into expected and unexpected rewards. Unexpected rewards had no significant effects; expected tangible rewards were found to significantly undermine both self-reported task interest and free-choice intrinsic motivation.

Using cognitive evaluation theory as their framework, Deci et al. (1999) further subdivided expected tangible rewards into task-noncontingent, engagement-contingent, completion-contingent, and performance-contingent rewards. Task-noncontingent rewards were ‘‘those given without specifically requiring the person to engage in the activity’’ (p. 636); engagement-contingent rewards were those offered to participants for engaging in a task without a requirement to complete the task, do it well, or reach some standard. Completion-contingent rewards were those offered and given for completing a task, and performance-contingent rewards were defined as those ‘‘offered dependent upon the participants’ level of performance’’ (p. 636). Deci et al. found no significant negative effects for task-noncontingent rewards; engagement-contingent rewards produced significant negative effects on both free-choice intrinsic motivation and self-reported task interest. Completion-contingent and performance-contingent rewards also resulted in significant negative effects on the free-choice intrinsic motivation measure.

In addition, Deci et al. (1999) provided a breakdown of performance-contingent rewards into studies of ‘‘maximum’’ and ‘‘not-maximum’’ reward. In studies of maximum reward, participants were offered rewards graded in terms of meeting a criterion or performance standard; all met the criterion and received the full amount of reward. Six studies were identified by Deci et al. as involving not-maximum reward. In these studies, some participants failed to attain the criterion and were given less than the maximum reward. Deci et al. reported that relative to a nonrewarded control condition, participants receiving less than the maximum reward showed a large decline in free-choice intrinsic motivation. In fact, the value ($d = -0.88$) was the largest mean effect size in their entire analysis.

As a supplemental analysis, Deci et al. (1999) analyzed studies with children in which the free-choice assessment of high-interest activities was conducted immediately following the removal of reward, within a week, and after a week. Deci et al. found negative effects at each time of assessment and suggested that the undermining effect is not a transitory phenomenon. An additional analysis of the effects of rewards on low-interest tasks was conducted by Deci et al.; no statistically significant effects were detected.

All in all, Deci et al.’s (1999) meta-analysis produced numerous negative effects of the various reward contingencies. Given the discrepancies between Deci et al.’s and Cameron and Pierce’s (1994) findings, it is important to examine carefully the procedures used by Deci et al. The first noteworthy difference between the two meta-analyses occurs at the level of all rewards. Cameron and Pierce were interested in assessing the overall effects of rewards across all types of tasks. Deci et al. did not conduct this analysis; instead, they argued that the more theoretically relevant question concerned the effects of rewards on tasks of high initial interest.

We contend that an analysis of the overall effect of reward is central to an understanding of this complex area of research. On a practical level, many educators, parents, and administrators have taken the position that rewards and incentive systems are harmful. The view is that rewards negatively affect students' intrinsic interest across all types of activities (e.g., reading, math, science, computer games, etc.); no distinction is made between low and high initial levels of task interest. Writers who caution against the use of rewards and reinforcement frequently use examples to illustrate their point. More often than not, activities such as reading, lawn mowing, and mathematics are cited as activities that people will lose interest in if they are given rewards for performing the activity. Most of these activities are not ones that individuals begin doing with high levels of initial interest. Importantly, policy makers who adopt the view that rewards are harmful rarely distinguish between high- and low-interest tasks. Because of this, an analysis of the overall effects of reward is warranted. It is our contention that a more complete hierarchical breakdown of the effects of rewards on intrinsic motivation should begin at the level of all rewards over all types of tasks. Following this, a breakdown of reward effects on high- and low-interest tasks would be appropriate.

A further difficulty with Deci et al.'s (1999) meta-analysis concerns their supplemental analysis of reward effects on low-interest tasks. Several studies that used low-interest tasks were excluded from their primary meta-analysis of high-interest tasks (e.g., Freedman & Phillips, 1985; Overskeid & Svartdal, 1996). The problem is that these studies were not brought back into their supplementary analysis of low-interest tasks.

Another concern is that for some studies in their analysis of high-interest tasks, Deci et al. (1999) omitted conditions that were relevant to their analyses. For example, in an experiment by Wilson (1978), one group was offered $0.50 to engage in the target activity, a second group was offered $2.50 and a control group performed the task without the offer of reward. In Deci et al.'s analyses, only one of the rewarded groups was included. For other studies that used more than one level of reward magnitude (e.g., Earn, 1982; McLoyd, 1979; Newman & Layton, 1984), Deci et al. included all reward conditions. Their omission of certain conditions within studies does not appear to be systematic (e.g., reward magnitude was not examined by Deci et al. as a moderator), yet there are a number of different types of cases in which this occurs. In addition, as did Cameron and Pierce (1994), Deci et al. also missed a few experiments that met their inclusion criteria and that were published during the period covered by their meta-analysis. Also, several studies using high-interest tasks that measured self-reported task interest were either excluded or inadvertently omitted from Deci et al.'s analyses. Many of these studies found positive effects on the self-report measure of task interest; Deci et al.'s omission of these effects helps to explain why they found either negative effects or no effects on the task-interest measure. A list of studies not included in Deci et al.'s analyses, dependent measures that were precluded, and a description of conditions omitted by Deci et al. are presented in Appendix A. Any computational differences in sample sizes and effect sizes are also outlined in Appendix A.

A final issue concerns the classification of studies into various reward contingencies. Deci et al. (1999) suggested that Cameron and Pierce (1994) miscategorized many experiments. Using cognitive evaluation theory to guide their classification of studies, Deci et al. established the categories of task-noncontingent, engagement-contingent, completion-contingent, and performance-contingent rewards. Although this categorization system may be informative for cognitive evaluation theory, the problem is that the categories are too broad. Studies that used very different procedures were pooled into overall categories of engagement-contingent, completion-contingent, and performance-contingent rewards. For example, under performance-contingent reward, Deci et al. pooled experiments in which participants were offered a reward for doing well, for each problem or unit solved, for obtaining a certain score, or for exceeding a norm. Eisenberger, Pierce, and Cameron (1999) examined some of these different reward procedures and found very different effects on measures of intrinsic motivation.

Rather than argue about which studies belong in which category, we suggest that a more nuanced approach is to return to the methods section of the original studies and code the reward procedures actually employed in the experiment. To the extent that one can obtain consistency in coding, such a procedural categorization of reward contingencies would allow an assessment of the effects of the actual contingencies rather than those presumed to be an effect by any theoretical perspective. The literature on rewards and intrinsic motivation is fraught with competing theories (e.g., cognitive evaluation theory, the overjustification hypothesis, social cognitive theory, general interest theory, the competing response hypothesis, behavioral theory). The problem with organizing studies according to a particular theoretical stance is that each theory could be used to organize the literature and, using categories appropriate to the theory, each theory could gain support. Using a theoretical approach to guide the classification of the reward procedures does not provide us with a definitive answer about the effects of reward contingencies on measures of intrinsic motivation. Instead, we propose that a procedural description of reward contingencies not only allows us to assess where we stand in terms of the effects of the actual reward contingencies but also provides us with a test of alternative accounts of the effects of rewards on intrinsic motivation.

A NEW META-ANALYSIS: RESOLVING DIFFERENCES BETWEEN PREVIOUS META-ANALYSES

To address criticisms and resolve discrepancies among Cameron and Pierce's (1994), Eisenberger and Cameron's (1996), and Deci et al.'s (1999) meta-analyses, we provide a reanalysis of the effects of rewards on intrinsic motivation. Our goal is to eliminate shortcomings of the prior meta-analyses (including ours) while building on their strengths. Our current meta-analysis focuses on how reward affects measures of free-choice intrinsic motivation and self-reported task interest. In accord with Deci et al., free-choice measures included free time on task when the rewards were removed and, when time measures were not available, performance during the free-choice period. As did Deci et al., we combined performance and time measures to make up the free-choice intrinsic motivation index (we found no significant differences in the analyses when only time measures were analyzed). Our analysis begins with an assessment of the overall effects of reward.

We then examine the effects of different moderator variables. To deal with Deci et al.'s (1999) criticism, our first breakdown is in terms of high and low initial task interest. Subdividing the studies by high and low initial task interest allows us to directly compare our findings with those of Deci et al. In doing so, we favor cognitive evaluation theory. On the other hand, failure to find pervasive negative effects even with high-interest tasks favors the conclusion that reward contingencies do not destroy interest. In other words, the strongest way to test Deci et al.'s claims is to use their requirement that tasks used in the studies must be broken out by high and low initial interest.

At each level of our analysis, a homogeneity statistic (Q) was calculated to determine whether the set of effect sizes could be considered homogeneous.

When Q was significant, we proceeded with further moderator analyses. Rewards were broken down by type (verbal or tangible), by expectancy (expected, unexpected), and by reward contingency. In addition, when there was enough data, we examined differences between studies in which participants received maximum or less than maximum rewards. These breakdowns generally resulted in homogenous samples. However, in a few cases, homogeneity could still not be obtained even after a thorough examination of potential moderators. In these cases, we conducted the analysis by removing outliers (as did Deci et al., 1999). Outliers were examined in an attempt to explain their extreme values. At each level of our analysis, we report mean effect sizes, significance tests, and 95% CI. However, we should point out that making conclusions based on heterogeneous samples may be misleading. In a hierarchical breakdown, interpretations should focus on the homogeneous effects at the bottom level of the analysis (see Hunter & Schmidt, 1990).

Sample of Studies

Studies included in the present analysis incorporated the databases of Cameron and Pierce (1994) and Deci et al. (1999). In addition, in a search of PsycINFO, we located a few new studies and a few studies that were inadvertently missed in previous meta-analyses. The criteria for including a study in a sample were as follows: A rewarded group was compared to a nonrewarded group, the rewards were distinguished as verbal (praise, positive feedback) or tangible (e.g., money, candy, good-player awards), and intrinsic motivation was measured as free choice (time spent on the task following the removal of reward or performance on the task during the free-choice period) or by self-reported measures of task interest (task liking, enjoyment, satisfaction, or task preference). Two studies included in Cameron and Pierce's research were omitted; in one study (Boggiano & Hertel, 1983), the dependent measure was assessed before all participants worked on the task; in the other study (Boal & Cummings, 1981), all participants (including the control group) received monetary payments. These studies were also not included in Deci et al.'s analyses.

In addition to including published work and in keeping with Deci et al. (1999), we included unpublished doctoral dissertations. The resulting sample consisted of 145 independent studies (21 of the experiments were from unpublished doctoral dissertations). Of these, 115 studies included a free-choice measure of intrinsic motivation; 100 included a self-report measure of task interest.

Classification and Coding of Studies

To code initial levels of task interest, we used the procedures described by Deci et al. (1999). If a measure of initial task interest was reported in the article, the study was classified as a low-interest task when the average on that measure was below the midpoint of the scale for the activity and as a high-interest task when the average was above the midpoint. Two studies not included in any of Deci et al.'s analyses (Freedman & Phillips, 1985; Phillips & Freedman, 1985) provided initial task-interest measures and included both a high- and a low-interest task. We included these studies in our analysis of low- as well as high-interest tasks (see comments in Appendix A). Studies without initial interest measures were classified as high or low depending on how the researcher defined the task or on whether the task had been described as interesting or uninteresting in prior experiments.

Studies were also classified according to reward type (tangible or verbal), reward expectancy (expected or unexpected), and reward contingency. To classify studies by reward contingency, we went back to the original studies, read the precise procedures used for re-

TABLE 1

Description of expected tangible reward contingencies

Reward contingency	Description
Task noncontingent	Reward is offered for agreeing to participate, for coming to the study, or for waiting for the experimenter. Offer of reward is unrelated to engaging in the task.
Rewards offered for doing well	Reward is offered for doing well on the task or for doing a good job. No specification is given as to what it means to do a good job or to do well.
Rewards offered for doing a task	Reward is offered to engage in the experimental activity. No instructions are given about how well participants must perform or whether they must complete the task.
Rewards offered for finishing or completing a task	Reward is offered to finish an activity, to complete a task, or to get to a ceratin point on the task. The reward is not related to quality of performance.
Rewards offered for each unit solved	Reward is offered for each unit, puzzle, problem, etc., that is solved.
Rewards offered for surpassing a score	Reward is offered for surpassing a particular specified score (absolute standard). In some cases, the better the score, the higher the reward.
Rewards offered for exceeding a norm	Reward is offered to meet or exceed the performance of others on the task (relative standard).

ward delivery, and wrote down what was said to participants and how the reward was delivered. We then organized the studies into seven main categories of reward contingency: rewards delivered regardless of task involvement (task noncontingent); rewards given for doing a task; rewards for doing well; rewards for finishing or completing a task; rewards given for each problem, puzzle, or unit solved; rewards for achieving or surpassing a specific score; and rewards for meeting or exceeding others. Although all studies were coded for reward contingency, it was at the level of expected (offered) tangible reward that it became necessary to analyze studies in the various reward contingencies. Other analyses resulted in homogeneity, and further breakdowns were not required. In Table 1, we provide definitions and descriptions for each of these contingencies at the level of expected tangible reward. A comparison of our reward contingencies and those of Deci et al. (1999) is presented in Appendix B. We return to a discussion of these comparisons in our results section.

In some studies, there was not enough information to code the contingency (e.g., Chung, 1995; Hom, 1987). In addition, a few studies used a contingency that did not fit into any of the seven categories; for example, W. E. Smith (1975) offered rewards to participants for showing signs of learning. These studies were included in overall analyses, but were omitted from the analysis of reward contingencies. A list of the studies used in each analysis, a description of reward type, reward expectancy, and reward contingency, together with effect sizes are presented in Appendixes C through G.

To ensure reliability of coding, the second author was given the definitions for each contingency (Table 1) and a sample of 32 studies to code (each of the studies involved expected tangible rewards). Reliability calculated as percentage agreement was 97% (31 of 32 studies). One study (L. W. Goldstein, 1977) included a condition in which participants were offered a reward to take pictures. The issue was whether this contingency involved reward simply for doing the task or for finishing the task. The third author was brought in to code the study; he pointed out that participants in the reward condition were not required to complete or finish the task to obtain the reward and that Goldstein stated that ‘‘the reward did not imply that the subject had done well on the task, only that s/he had engaged in it’’ (p. 30). Hence, the reward contingency was classified as a reward offered for doing the task.

Finally, we identified studies that involved maximum or less than maximum reward. Such studies involved offering participants a reward for doing well, for finishing a task, for each problem or unit solved, for surpassing a score, or for exceeding a norm. Studies were considered to produce the maximum reward if participants in the reward condition met the performance requirements and received the full reward. Less than maximum reward occurred when there was a time limit such that some participants were unable to meet all the requirements in the time allotted and were given less than the full reward. For example, Deci’s (1971) experiment involved less than maximum reward. Participants were offered $1.00 for each of four puzzles solved within a 13-min time limit. Not all participants were able to solve the puzzles within the time limit and did not receive the full reward.

Calculation and Analysis of Effect Sizes

After all studies were coded, we calculated effect sizes (g) for each comparison of a rewarded group to a nonrewarded group on the free-choice and self-report measures of intrinsic motivation. Positive effect sizes indicate that rewards produced an increase in measures of intrinsic motivation relative to a control group, negative effect sizes denote a decrease, and an effect size of 0.00 indicates no difference. When there was not enough information to calculate an effect size, we attempted to contact the researchers. From a list of 22 researchers, we were able to locate E-mail addresses for nine. E-mail messages were sent requesting the missing data. Eight people replied; six could not locate the data, and two provided us with data for studies by Wicker, Brown, Wiehe, and Shim (1990) and by Dollinger and Thelen (1978). When we could not obtain missing data, we imputed an effect size of 0.00. Each analysis was conducted with zeros included and excluded. In accord with Deci et al. (1999), we report the analyses with the zeros included; however, when mean effect sizes were altered to any extent by the inclusion of zeros, we report the analysis with and without zeros.

Eisenberger, Pierce, and Cameron (1999) pointed out that there were two possible types of control comparisons for some of the studies labeled performance contingent by Deci et al. (1999). In some studies, the control group was told the performance objectives and was given performance feedback (complete control); in others, the control group was not told a performance objective and no feedback was given (partial control). Eisenberger, Pierce, and Cameron examined differences between these two types of comparisons (reward vs. partial control, reward vs. complete control). One small difference was detected on the free-choice measure. When rewards were offered to exceed others, reward versus a partial-control condition resulted in a nonsignificant positive effect; the mean effect for reward versus a complete control was significantly positive (no other comparisons resulted in differences). Because this difference was small and both mean effects were in the same direction, we included studies with either type of control condition in the present analyses. If a study contained both types of controls (e.g., Harackiewicz, Manderlink, & Sansone, 1984), one effect size was calculated comparing the reward condition to both controls.

In accord with Deci et al. (1999) and with our previous procedures, more than one effect size was calculated for several studies in our analyses. For example, if a single study assessed free choice and used two types of expected tangible rewards (e.g., rewards offered

TABLE 2

Hierarchical analysis of the effects of rewards on measures of intrinsic motivation

Analysis of the effects of reward	*K*	*N*	*d*+	95% CI
Free-choice intrinsic motivation				
All reward[a]	115	8,176	−0.08	−0.12, 0.02
Low initial task interest	12	429	0.28*	0.07, 0.47
High initial task interest[a]	114	7,888	−0.09*	−0.14, −0.04
Verbal reward	25	1,374	0.31*	0.20, 0.41
Tangible reward[a]	102	6,942	−0.17*	−0.22, −0.12
Unexpected reward	9	375	0.02	−0.18, 0.22
Expected reward (offered)[a]	101	6,703	−0.18*	−0.23, −0.13
Self-reported task interest				
All reward[a]	100	8,028	0.12*	0.07, 0.16
Low initial task interest	11	503	0.12	−0.06, 0.30
High initial task interest[a]	98	7,547	0.12*	0.07, 0.17
Verbal reward[a]	24	1,584	0.32*	0.22, 0.43
Tangible reward[a]	83	6,354	0.08*	0.03, 0.13
Unexpected reward	5	299	0.03	−0.20, 0.26
Expected reward (offered)[a]	81	6,138	0.08*	0.03, 0.13

Note. K = number of studies; N = total sample size; $d+$ = mean weighted effect size; 95% CI = 95% confidence interval.
[a] The sample of effect sizes was significantly heterogeneous.
* $p < .05$.

for doing the task and rewards offered for surpassing a certain score) plus a control group, two effect sizes were calculated. Each individual effect size was entered into the relevant analysis (expected tangible rewards for doing a task, expected tangible rewards for surpassing a score). For the analyses of expected tangible reward, tangible reward, and all reward, one effect size was calculated (the two groups were compared to the control group) and entered into the overall analyses. This strategy satisfies the independence assumption of meta-analytic statistics (Hedges & Olkin, 1985) and gives equal weight to each study analyzed. Thus, subcategories (e.g., rewards offered for doing the task, for doing well, etc.) may contain more effect sizes than the superordinate category (expected tangible reward). For example, for all reward on the free-choice measure (over both high- and low-interest tasks), there were 126 effect sizes, but only 115 of these are independent (several are within the same study).

After all effect sizes were calculated, the present analyses were run on the computer program Meta (Schwarzer, 1991) using the weighted integration method described in our section on meta-analytic procedures. The program converts effect size, g, to d; mean weighted effect size ($d+$) is obtained; 95% CI is constructed around the means, and a homogeneity statistic, Q, is computed.

RESULTS OF OUR META-ANALYSIS

In Table 2, we present the results for our meta-analysis up to the level of reward contingency. Table 2 presents mean weighted effect size ($d+$) and 95% CI for each analysis. Mean effects are considered statistically significant when the CI does not include zero. In the present meta-analysis, positive effect sizes indicate that reward produces increases in intrinsic motivation, negative effect sizes support the claim that rewards undermine intrinsic motivation, and zero effects indicate no evidence for an effect of reward. According to J. Cohen

(1988), an effect size of ±0.20 is considered small, ±0.50 is moderate, and greater than ±0.80 is large.

All Rewards

First, the overall effects of reward were analyzed across all conditions and across high- and low-interest tasks. On the free-choice measure, Table 2 indicates that there was no significant effect ($d+ = -0.08$, CI = −0.12, 0.02). On the measure of self-reported task interest, a small significant positive effect was detected ($d+ = 0.12$, CI = 0.07, 0.16). This analysis was not conducted by Deci et al. (1999); therefore, the findings cannot be compared. The results are, however, in accord with those of Cameron and Pierce (1994). On both the free-choice and self-report measures, however, the sets of studies were significantly heterogeneous, suggesting the necessity of a moderator analysis. Thus, at the next level of analysis, we divided studies into those with low- and high-interest tasks.

The Effects of Rewards on Low-Interest Tasks

When reward effects were analyzed for tasks with low initial interest, Table 2 shows a statistically significant positive effect on the free-choice measure ($d+ = 0.28$, CI = 0.07, 0.47); there was no significant effect on self-reported task interest ($d+ = 0.12$, CI = −0.06, 0.30). These findings indicate that when a task is not initially interesting, rewards enhance free-choice intrinsic motivation but not verbal expressions of task interest.

Although the studies in this analysis were considered homogeneous (i.e., Q was not significant), we examined whether there were any differences among different types of rewards, expectancies, and contingencies. On the free-choice measure, only one study included a condition that used a verbal reward (the effect was positive). For tangible reward, one study included an unexpected reward condition (the effect was positive). All of the 12 studies with low-interest tasks included an expected tangible reward condition; compared with a nonreward control, the mean effect was significantly positive ($d+ = 0.26$, CI = 0.06, 0.45). Nine studies involved offering the reward for doing the task; on the free-choice measure the effect remained significant ($d+ = 0.26$, CI = 0.03, 0.48). For self-reported task interest, no significant effects were found under any of the conditions.

In Deci et al.'s (1999) supplemental analysis of low-interest tasks (p. 651), fewer studies were included and no significant effects were found on either the free-choice or the self-report measures of intrinsic motivation.

The Effects of Rewards on High-Interest Tasks

For high-interest tasks, the mean effect size on free choice (Table 2) showed a small but significant negative effect ($d+ = -0.09$, CI = −0.14, −0.04); the set of effect sizes, however, was heterogeneous. The mean effect size for self-reported task interest was significant, small, but in a positive direction ($d+ = 0.12$, CI = 0.07, 0.17); the sample of effect sizes was also heterogeneous. Deci et al. (1999) also reported a significant negative effect on the free-choice measure but a nonsignificant effect on the task-interest measure. As noted, Deci et al. omitted or missed several self-report effect sizes.

Verbal Rewards

Verbal rewards were found to significantly enhance both free-choice intrinsic motivation ($d+ = 0.31$, CI = 0.20, 0.41) and self-reported task interest ($d+ = 0.32$, CI = 0.22, 0.43). These results were also obtained by Deci et al. (1999), who reported similar small to moderate positive effects of verbal rewards.

On the free-choice measure, the set of effect sizes was homogeneous, suggesting that no further breakdowns were necessary. In most studies of verbal reward, the rewards were unexpected and the mean effect was positive; a positive effect was also found with the five studies that used expected rewards. In addition, verbal rewards were generally delivered simply for doing a task and were not contingent on any specific level of performance (again, the effects were positive). When the effects of verbal reward on free choice were examined with children versus adults (mainly college students), children showed a smaller positive effect ($K = 10$, $N = 320$, $d+ = 0.22$, CI = 0.04, 0.39) than adults ($K = 15$, $N = 844$, $d+ = 0.36$, CI = 0.22, 0.49). Deci et al. (1999) also reported a larger effect for adults but a nonsignificant effect for children (our effect size for children was statistically significant because we included more studies than Deci et al.).

On the task-interest measure, the set of effect sizes for verbal reward was significantly heterogeneous. We conducted moderator analyses of children versus adults and expected versus unexpected reward. Mean effect sizes for each of these analyses remained significantly positive, but homogeneity was still not obtained. In almost all studies, the rewards were given for doing the task; hence, this reward contingency could not be a moderator.

To obtain homogeneity, three studies were removed from the analysis (the same outliers were removed by Deci et al., 1999). Inspection of the outliers indicated that two of the studies (Butler, 1987; Vallerand, 1983) produced large positive effects; these studies did not differ in obvious ways from other studies in the sample except for their tendency to generate extreme values of effect size. The third outlier (Kast & Connor, 1988) produced a negative effect (−0.46). Kast and Connor compared control participants to participants who were praised for their performance on the task as well as to another group who were also praised but who were told that they should be doing well. The second verbal reward condition produced a negative effect and was different from verbal reward used in other studies; Deci et al. termed this ‘‘controlling’’ reward. When the outliers were removed from the analysis of verbal rewards on the task-interest measure, the set of studies was homogeneous and the mean effect remained significantly positive ($K = 21$, $N = 1{,}194$, $d+ = 0.32$, CI = 0.21, 0.44). In this data set, there were six studies that did not provide enough information to obtain an estimate of effect size (these studies were given an effect size of 0.00). When these studies were removed, the mean effect size for task interest showed a slight increase ($K = 15$, $N = 981$, $d+ = 0.40$, CI = 0.27, 0.53).

Tangible Rewards

When the effects of tangible rewards on high-interest tasks were analyzed, Table 2 shows a small significant negative effect on the free-choice measure ($d+ = -0.17$, CI = −0.22, −0.12) and a small significant positive effect on self-reported task interest ($d+ = 0.08$, CI = 0.03, 0.13). Both of these samples of effect sizes were significantly heterogeneous and required a further moderator analysis.

Reward expectancy. Tangible rewards were subdivided into unexpected (rewards delivered without a statement of the contingency) and expected (rewards delivered after a statement of contingency) categories. No significant effects were detected for unexpected tangible rewards (see Table 2), and the samples were homogenous (Deci et al., 1999, reported similar findings). Expected tangible rewards produced a negative effect on the free-choice measure ($d+ = -0.18$, CI = −0.23, −0.13) and a positive effect on the self-report measure ($d+ = 0.08$, CI = 0.03, 0.13), but both of these samples were significantly heterogeneous.

Reward contingency. For the next level of analysis, expected tangible rewards were subdivided into various reward contingencies. Results of our analysis on the free-choice measure are presented in Figure 1. No significant

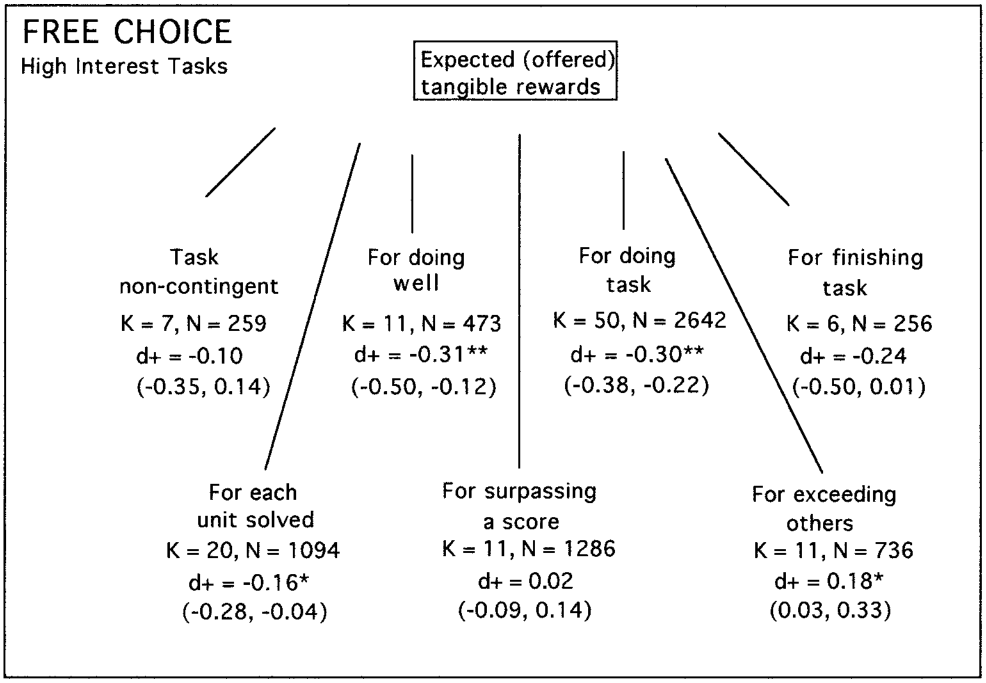

Figure 1. The effects of expected tangible reward contingencies on free-choice intrinsic motivation under high levels of initial task interest. K = number of studies, N = total sample size, $d+$ = mean weighted effect size; statistically reliable effect sizes are marked with an asterisk (*$p < .05$, **$p < .01$). Positive effect sizes indicate higher intrinsic motivation for rewarded versus control groups; negative effect sizes indicate lower intrinsic motivation for rewarded groups. Numbers in parentheses represent 95% confidence intervals. All effect sizes are based on homogeneous samples.

effects were detected when the rewards were task noncontingent, were offered for finishing or completing a task, or were offered for attaining or surpassing a score. Figure 1 shows significant negative effects when the rewards were offered for doing a task, for doing well on a task, and for each unit solved. A significant positive effect was found when the rewards were offered for meeting or exceeding the performance level of others.

When rewards were offered for doing a task, the effect was significantly negative (K = 57, N = 2,910, $d+ = -0.35$, CI = -0.43, -0.27) but not homogeneous. Although we searched for moderators (salient vs. nonsalient reward, children vs. adults, and time of reward delivery), analyses of these variables did not result in homogeneous samples. As a result, outliers were identified and omitted. The mean effect with outliers removed is presented in Figure 1. Two of the outliers produced positive effects; the only differences between these two studies and the bulk of studies were that the study by Tripathi and Agarwal (1988) was conducted in India and the study by Brennan and Glover (1980) was designed to assess the effects of rewards when the rewards were shown to function as reinforcement. Other outliers (Chung, 1995; Danner & Lonkey, 1981; Fabes, Eisenberg, Fultz, & Miller, 1988; Morgan, 1983, Experiment 1; Okano, 1981, Experiment 2) had large negative effects but there was no common factor that could explain their extreme values.

Our findings for free choice indicate that when reward contingency is defined in terms of experimental procedures, negative, neutral, and positive effects are obtained. Using cognitive evaluation theory as their framework for the categorization of reward contingencies, Deci et al. (1999) found neg-

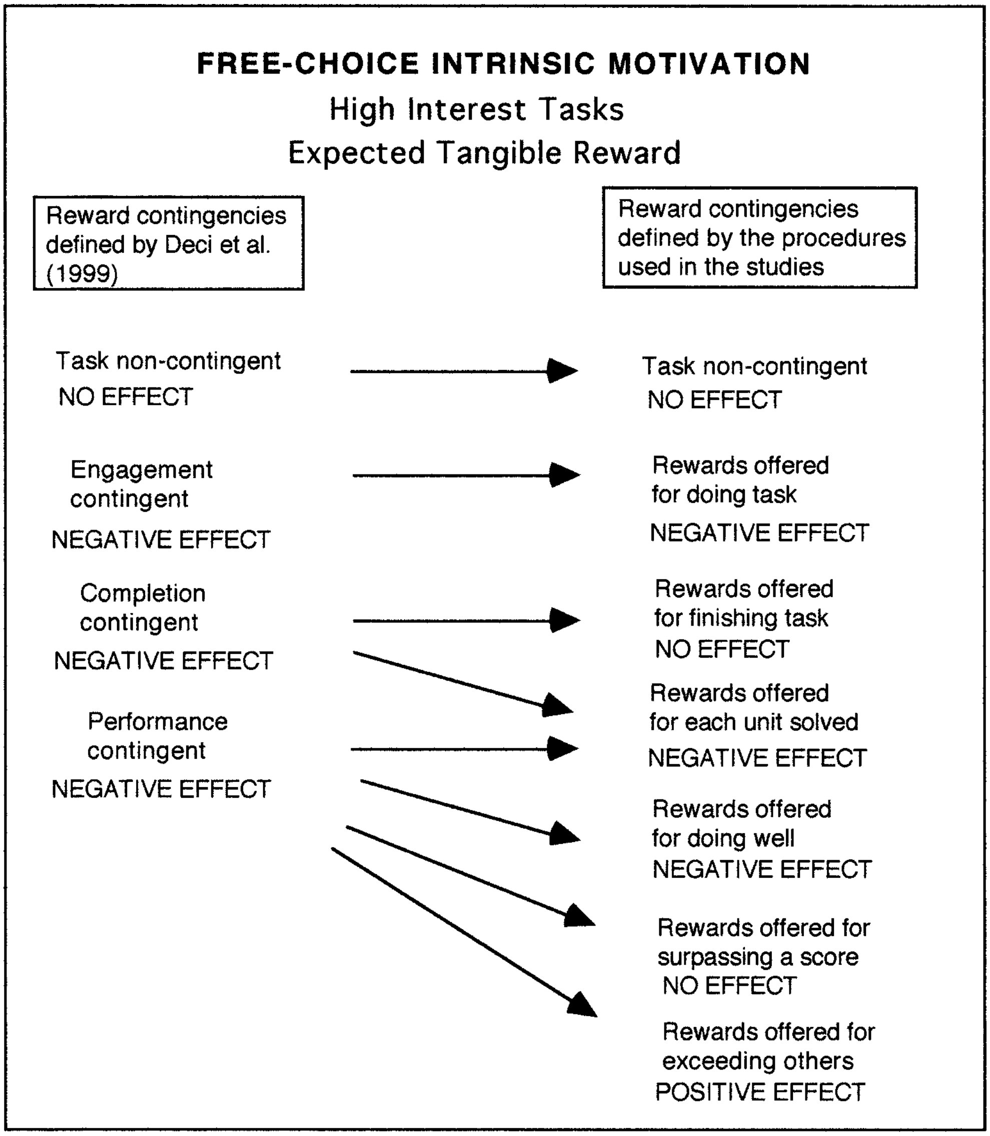

Figure 2. A comparison of our findings with Deci et al.'s (1999) effects of expected tangible reward contingencies on free-choice intrinsic motivation for high-interest tasks. Deci et al.'s categories of completion-contingent and performance-contingent reward contained studies that involved "reward offered for each unit solved."

ative effects for all but task-noncontingent rewards. One way to understand these differences is to compare Deci et al.'s effects and definitions of contingencies with our effects and procedural definitions. Figure 2 shows this comparison and indicates that Deci et al.'s completion-contingent and performance-contingent rewards consisted of a variety of reward procedures having different effects.

Our results for the task-interest data are presented in Figure 3. The analysis shows no significant effect for task-noncontingent rewards, a small significant negative effect for rewards offered for doing, and significant positive effects for each of the other contingen-

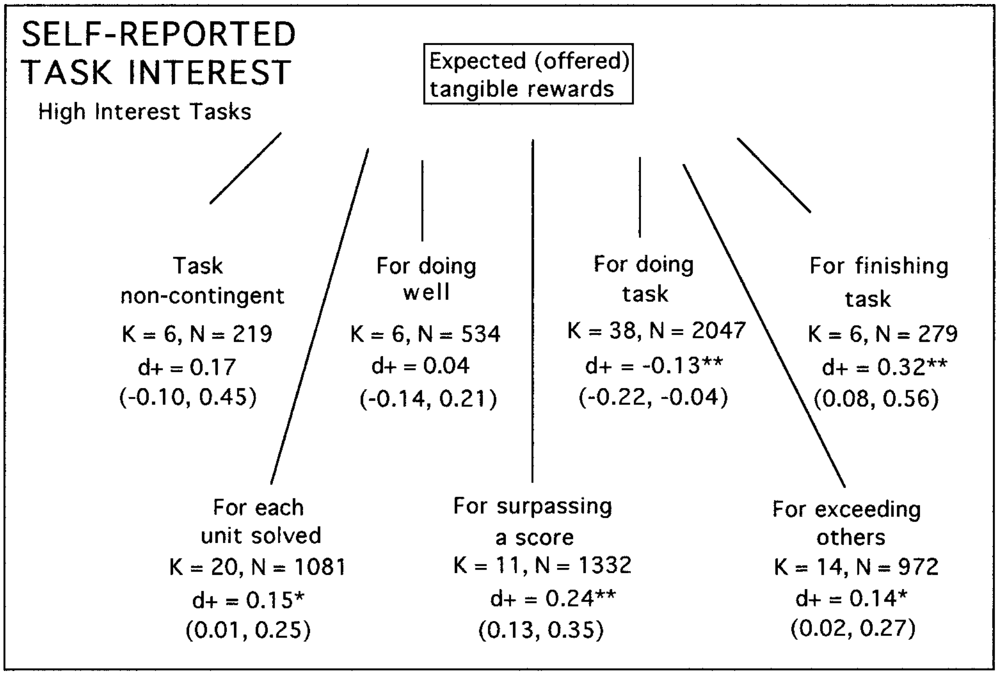

Figure 3. The effects of expected tangible reward contingencies on self-reports of task interest under high levels of initial task interest. K = number of studies, N = total sample size, $d+$ = mean weighted effect size; statistically reliable effect sizes are marked with an asterisk (*$p < .05$, **$p < .01$). Positive effect sizes indicate higher intrinsic motivation for rewarded versus control groups; negative effect sizes indicate lower intrinsic motivation for rewarded groups. Numbers in parentheses represent 95% confidence intervals. All effect sizes are based on homogeneous samples.

cies. In the analysis of rewards offered for doing, 14 studies were given effect sizes of 0.00; when these studies are removed from the analysis, the negative effect increased from −0.13 to −0.22 (K = 24, N = 1,201, $d+$ = −0.22, CI = −0.33, −0.10).

In the analysis of rewards offered for each unit completed, when all studies were included the effect was positive (K = 22, N = 1,161, $d+$ = 0.19, CI = 0.08, 0.31) but significantly heterogeneous. Two studies (Kruglanski et al., 1975, Experiment 1; Wimperis & Farr, 1979) had a large positive effect size; when these studies were omitted, homogeneity was attained (Figure 3 presents the data for homogenous samples).

In Figure 4, we compare Deci et al.'s (1999) findings and reward contingencies with ours. For completion-contingent and performance-contingent rewards, Deci et al. found no significant effects, whereas our findings show a number of positive effects for studies that would be included in these categories. As discussed previously, many studies with self-report measures were not included in Deci et al.'s analyses (see details in Appendix A).

Maximum versus less than maximum reward. On the free-choice measure of intrinsic motivation, there was only one reward contingency (rewards offered per unit solved) that allowed a comparison between maximum and less than maximum reward. For other reward contingencies, most studies involved maximum reward; a comparison with less than maximum reward would be unreliable. When rewards were offered for each unit solved, the findings showed nonsignificant effects for studies of maximum rewards (K = 6, N = 345, $d+$ = −0.03, CI = −0.25, 0.18) and a significant negative effect for studies of less than maximum reward (K = 14, N = 749, $d+$ = −0.22,

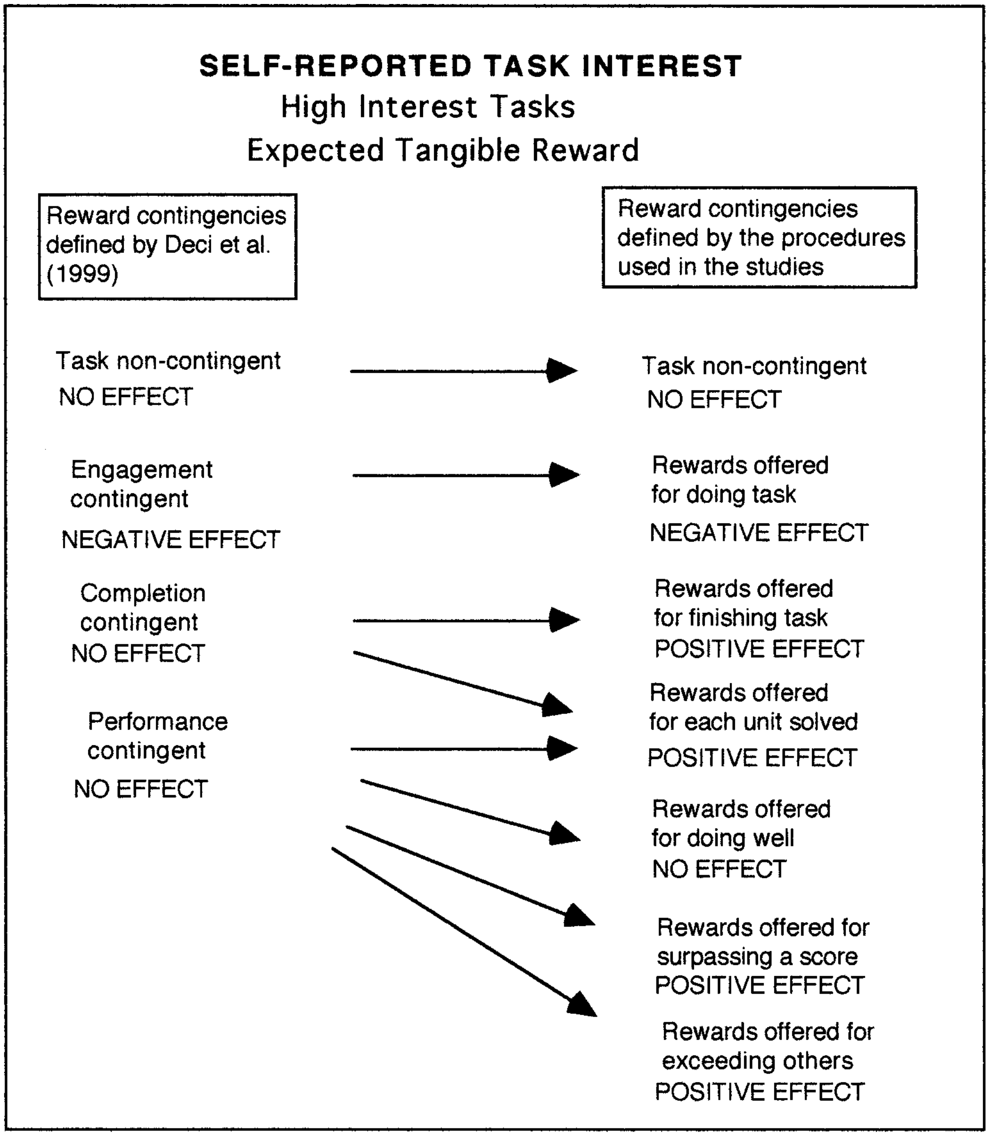

Figure 4. A comparison of our findings with Deci et al.'s (1999) effects of expected tangible reward contingencies on self-reports of task interest for high-interest tasks. Deci et al.'s categories of completion-contingent and performance-contingent reward contained studies that involved "reward offered for each unit solved."

CI = −0.37, −0.07). These two sets of effect sizes were homogeneous. These results suggest that the negative effect of pay per unit is associated with participants receiving less than maximum rewards.

No analyses were conducted on differences between maximum and less than maximum rewards on the self-report measure. Most of the contingencies had too few studies that used less than maximum reward. For studies involving the offer of reward for each problem solved, there were too few experiments of maximum reward (see Appendix F).

DISCUSSION

A major issue in psychology and education is that rewards and reinforcement have a detrimental effect on intrinsic motivation. The concern is that if people receive reinforcement or rewards for activities they already enjoy, they will be less motivated to engage in those activities than they were prior to the introduction of reward once the rewards are no longer forthcoming. In other words, rewards and reinforcement are said to decrease intrinsic motivation. Since the 1970s, over 100 studies have been conducted to assess the effects of reward on intrinsic motivation. The vast majority of studies on the topic have employed between-groups statistical designs. Rewarded participants are compared to nonrewarded controls. Intrinsic motivation is measured by the difference between groups on task interest and free choice (time on task and performance on task once the rewards are removed). A meta-analysis of this experimental literature by Cameron and Pierce (1994) and Eisenberger and Cameron (1996) found limited negative effects of rewards, whereas a more recent analysis by Deci et al. (1999) showed pervasive negative effects. The meta-analysis presented in this article was designed to correct flaws in the previous reviews and to resolve differences.

A Summary of Our Findings

In terms of the overall effects of reward, our meta-analysis indicates no evidence for detrimental effects of reward on measures of intrinsic motivation. This finding is important because many researchers and writers espouse the view that rewards, in general, reduce motivation and performance. In addition, many students of psychology and education are taught that rewards are harmful and that reward procedures should be avoided in applied settings. Our finding of no overall effect of reward, however, must be treated with caution. In our meta-analysis, the overall reward category lacked homogeneity, indicating the appropriateness of a moderator analysis. In other words, the overall reward category is too inclusive; rewards have different effects under different moderating conditions.

Figure 5 shows the effects of different moderating conditions. The effects of rewards on free-choice intrinsic motivation and self-reported task interest are presented only for homogeneous subsets. When a result was heterogeneous, we broke down the subset of effect sizes by different moderator variables until homogeneity was attained. A positive effect indicates that rewards enhanced the measure of intrinsic motivation relative to a control condition, a negative effect indicates a decrease for the rewarded group, and a zero effect indicates no significant effect.

The effects of all rewards are first broken into high- and low-interest tasks. When the tasks used in the studies are of low initial interest, rewards increase free-choice intrinsic motivation and leave task interest unaffected. This finding indicates that rewards can be used to enhance time and performance on tasks that initially hold little enjoyment. As Bandura (1986) recognized, ''Most of the things people enjoy doing for their own sake had little or no interest for them originally. . . . But with appropriate learning experiences, almost any activity . . . can be imbued with consuming significance'' (p. 241). Our results suggest that reward procedures are one way to cultivate interest in an activity. In education, a major goal is to instill motivation and enjoyment of academic activities. Many academic activities are not of high initial interest to students. An implication of our finding is that rewards can be used to increase performance on low-interest academic activities.

For high-interest tasks, verbal rewards are found to increase free choice and task interest. This finding replicates the results of Cameron and Pierce (1994) and Deci et al. (1999). Most social interaction in business, education, and clinical settings involves the use of verbal praise and positive feedback from managers, teachers, and therapists. When praise and other forms of

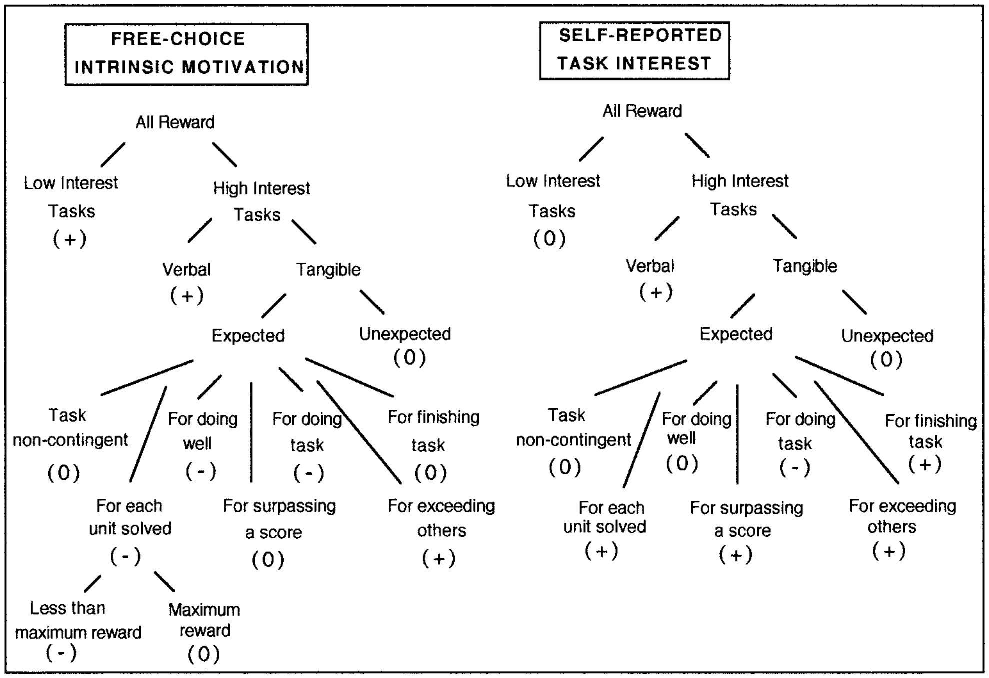

Figure 5. A summary of the meta-analysis comparing free-choice intrinsic motivation and self-reported task interest. 0 = no reliable effect; − = statistically significant negative effect of reward; + = statistically significant positive effect of reward.

positive feedback are given and later removed, our findings indicate that interest and performance increase.

The effects of tangible reward on measures of intrinsic motivation differ by reward expectancy. When rewards are delivered unexpectedly (without a description of the reward contingency), there is no evidence of a significant effect on either free choice or task interest. This suggests that it is not tangible rewards per se that undermine motivation and interest; instead it depends on instruction and the statement of contingency.

For high-interest tasks, when the rewards are tangible and expected (offered beforehand), there are different effects depending on the description of the reward contingency. When the offer of reward is unrelated to task behavior (task noncontingent), there is no evidence for an effect of reward on either free choice or task interest. On the other hand, when people are offered a tangible reward for doing a task or for doing well at a task, they often choose to do the activity less in a free-choice period. The negative effect of rewards offered for doing a task is also detected for the task-interest measure. We did not find a negative effect on task interest when the rewards are offered for doing well. One possibility is that the true effect is negative but, at this point, there are too few studies to yield a reliable estimate. In general, when the description of the reward contingency implies that rewards are loosely tied to performance, the evidence suggests that people show a small reduction in performance and interest.

Figure 5 shows that rewards offered for finishing or completing a task have a nonsignificant effect on the free-choice measure but a positive effect on task interest. Again, there were few studies in this category, and a firm conclusion about the effects is premature. Stronger conclusions can be drawn for the analysis of rewards offered for each unit solved. When participants are offered a reward for each problem, puzzle, or unit solved, our findings indicate a negative effect on free choice and a positive effect on task interest. A supplementary analysis involving less than maximum reward and maximum reward shows that the negative effect on free choice occurs when participants obtain less than the full reward. In studies of less than maximum reward, participants are given a time limit to solve problems. Thus, the negative effect may be a result of time pressure rather than reward. Another way to understand this result is to consider what less than maximum reward signifies to participants. If people are told they can obtain a certain level of reward but are given less than that level, they have received feedback information that indicates failure. In other words, this type of situation may represent failure feedback, not reward. When participants are not under time pressure and are able to obtain the maximum reward, there is no significant effect on the free-choice measure.

When rewards are offered for meeting or surpassing a score, Figure 5 shows no significant effect on free choice but a significant positive effect on task interest. Rewards offered for attaining a criterion are tightly linked to level of performance. In this situation, the rewards are tied to challenge and mastery of the activity, and people express interest in the task (see Bandura, 1986). When rewards are given for exceeding the performance level of others, the results show a significant increase on free choice and task interest. One possible explanation for the positive effects of this type of reward contingency is that rewards signify competence, self-efficacy, or ability at the task, and people enjoy doing activities that reflect their competence.

Overall, our analysis shows that tangible rewards can be used to produce both negative and positive effects on measures of intrinsic motivation. Positive effects are obtained when the rewards are explicitly tied to performance standards and to success; negative effects are produced when rewards signify failure or are loosely tied to behavior.

Durability of Reward Effects

Deci et al. (1999) have claimed that negative effects of rewards are not temporary. In a supplemental analysis, Deci et al. examined studies of children in which the free-choice assessment was conducted within a week following the removal of reward and after a week. Their analysis showed negative effects on free choice for each time of assessment. Deci et al. concluded that their results ‘‘indicate quite clearly that the phenomenon of extrinsic rewards undermining intrinsic motivation is not merely transitory’’ (p. 650). An examination of the studies included in Deci et al.’s supplementary analysis indicates that most of the effect sizes were based on rewards offered for doing the task or for doing well. When the free-choice assessment was conducted within a week following the removal of reward, 10 of 12 studies involved rewards offered for doing well or for doing the task. Of the 14 studies with assessments conducted more than a week later, 13 were concerned with rewards offered for doing well. Our interpretation of Deci et al.’s findings is that it is rewards offered for doing (or doing well) that continue to produce a negative effect on free choice, not extrinsic rewards in general. According to Bandura (1986), this kind of reward procedure imparts little indication of competence, in that the rewards are allocated without regard to quality of performance and are thus loosely tied to behavior.

An unresolved issue is whether there is a change in free-choice intrinsic motivation over time. We examined seven between-groups design studies of rewards offered for doing the task that assessed whether negative effects were maintained over time (Chung, 1995; Loveland & Olley, 1979; Morgan, 1983, Experiments 1 and 2; Ogilvie & Prior, 1982; Ross, 1975, Experiment 1;

Shiffman-Kauffman, 1990). These studies included two measures of free-choice intrinsic motivation, one after the removal of reward and a second a few weeks later. Only two of the seven studies showed a significant negative effect on the second measure (Morgan, 1983, Experiments 1 and 2). These results suggest that rewards offered for doing a task have transitory effects when multiple measures of free-choice motivation are used. This conclusion is strengthened by examining the results from studies using repeated presentations of reward followed by repeated assessments of intrinsic motivation following the removal of reward.

As previously indicated, some operant researchers tested the effects of rewards on intrinsic motivation by experimental designs in which the same individual was exposed to a baseline period, a reward intervention, and a return to baseline (Davidson & Bucher, 1978; Feingold & Mahoney, 1975; Mawhinney et al., 1989; Skaggs et al., 1992; Vasta et al., 1978). Participants were measured repeatedly during each phase of the experiment, and rewards were shown to increase measures of performance, indicating that the rewards functioned as reinforcement. The results of these experiments showed that participants spent as much (or more) time on the target activity in the postreward phase as they did in the initial baseline period. These findings indicate that negative effects of reward do not persist when task performance is rewarded on repeated occasions.

Magnitude and Impact of Reward Effects

It may be informative to consider how serious the negative effects are on high-interest tasks when the rewards are tangible, expected, and given for doing a task or for doing well on a task. In all of the studies involving these contingencies, time spent on the task during the free-choice period was the measure of free-choice intrinsic motivation. Using the free-time measure, one could ask how much less time students would spend on high-interest tasks (e.g., art, music, reading, drama) if a teacher implemented a reward system for doing the task (or doing well) and then removed it. Results from our meta-analysis indicate that the average effect size for a comparison between students who receive this reward procedure and nonrewarded individuals on time on task is about −0.30.

In the original experiments, free time on task was typically measured over an 8-min period. To convert the effect size of −0.30 to real time, one needs to know the pooled standard deviation of rewarded and nonrewarded groups. Because many researchers report only *t* or *F* statistics that cannot be converted to the overall pooled standard deviation, we are unable to provide an estimate of this parameter. Instead, we will use a well-designed study by Pretty and Seligman (1984) that provides a pooled standard deviation. Pretty and Seligman conducted two experiments with large samples and readily available statistical information. Both experiments compared a condition of tangible rewards offered for doing a high-interest task (Soma puzzles) with a nonrewarded control group on 8 min of free time. The pooled standard deviation was 2.6 min (Deci, 1971, also used Soma, the free-time measure was assessed over an 8-min period, and the pooled standard deviation was 2.4 min).

Using 2.6 min as the estimate of error, we are able to convert the negative effect size from the meta-analysis into real time. An effect size of −0.30 would mean than in an 8-min period, the average individual who is offered a tangible reward for doing the task (or doing well) will spend about 47 s less on the task when the reward is withdrawn than the average nonrewarded individual. Given this result, what would happen if a teacher implemented this incentive procedure in a reading program (for children who already enjoy reading) and then removed it? According to this estimate, students who are offered gold stars for reading would spend about 4 min less reading in a 40-min free-choice period than students not given the incentive. If we assume that students without reward spend about 30 min reading in the 40-min free-choice period, then rewarded students would spend about 26 min reading (based on Deci et al.'s, 1999, analysis of engagement-contingent reward, rewarded children would spend about 25 min reading). A 4- to 5-min reduction in free-time reading could be behaviorally important if cumulated over many successive opportunities to read, but there are no studies that have addressed this issue.

A cautionary note is in order. Our example of reading and reward depends on the use of a standard deviation from a single well-designed study. It also depends on the ability to extrapolate from an 8-min experimental period to longer ones. It is possible that the negative effects, such as they are, are evident only for a short time at the beginning of the free-choice period. That is, it may well be the case that if an hour of free choice were given, results might look very different. The point is that this is a hypothetical example. Further evidence is required to generalize the findings to experiments with longer free-choice periods or to everyday settings in which choice is distributed over long periods of time.

Given the state of the literature, we conclude that the negative effect of tangible rewards offered for doing a high-interest task (or doing well) is statistically significant, but the size of the effect does not suggest a strong impact. Of course, our conclusion with regard to the magnitude of the negative effects of reward contingencies applies equally to positive effects. That is, although the positive effects are statistically significant, they too are small.

A Comparison of Our Findings to Those of Deci et al. (1999)

Our pattern of findings for expected tangible reward contingencies differs from the results of Deci et al.'s (1999) meta-analysis. Deci et al. present a picture of pervasive negative effects. The picture depicted in our analysis is one of circumscribed negative effects. As noted, Deci et al. used reward contingencies that were theoretically relevant, but that were collapsed over distinct reward procedures. For example, on free-choice intrinsic motivation, Deci et al. showed a negative effect for performance-contingent rewards. The performance-contingent category included some studies of rewards offered for each unit solved, rewards offered for doing well, rewards offered for surpassing a score, and rewards offered for exceeding others. By combining these distinct procedures, Deci et al. obtained an overall negative effect for performance-contingent reward. We show that these diverse reward procedures produce different effects on free choice; hence, it is unwise to collapse them into a single category of performance-contingent reward. Similarly, Deci et al. collapsed over reward categories for the task-interest measure, and similar problems arise. In addition to collapsing over different reward categories, Deci et al. omitted several positive effects that, when included, resulted in positive findings for task interest. Overall, our meta-analysis indicates that rewards do not have pervasive negative effects when minor improvements to Deci et al.'s categorization of reward contingencies are made and all available studies are included.

Using cognitive evaluation theory to guide the classification of studies, Deci et al. (1999) obtained negative effects of tangible reward contingencies. We showed that by classifying studies according to the actual contingency used, different effects were obtained. That the results of a meta-analysis can be drastically altered by assigning studies to categories based on a particular theoretical orientation points to some important issues and limitations in this literature. The difference between our findings and those of Deci et al. points to a lack of standardization of reward procedures and definitions and suggests that, overall, the literature on rewards and intrinsic motivation is one of meager effects. One implication is that extreme caution must be exercised before making any applied policy decisions based on this body of research.

Theoretical Implications of Our Meta-Analysis

In terms of theoretical considerations, results from our reanalysis can be well explained by theories that predict that the effects of reward on intrinsic motivation depend on a clear specification of the reward contingency. For example, social cognitive theory (Bandura, 1986) predicts that rewards tied to level of performance enhance self-efficacy to the extent that the person is able to attain the performance standard (i.e., succeed). Greater self-efficacy leads to higher interest in a task and to more time spent on the activity. In our analyses, the positive effects of rewards given for surpassing a score or exceeding others are in accord with this account. The results also support a behavioral approach (e.g., Dickinson, 1989) in the sense that rewards that are closely tied to performance lead to interest in an activity; there is no evidence of negative effects on time spent on the activity when the rewards are withdrawn (participants return to baseline levels of the activity).

Our pattern of findings contradicts a strict application of cognitive evaluation and overjustification theories. Cognitive evaluation theory emphasizes the controlling aspect of performance-contingent rewards in reducing personal autonomy or self-determination. The loss of perceived autonomy leads to a loss of intrinsic motivation. Overjustification theory emphasizes the shift in attribution from internal to external sources that performance-contingent rewards produce. Both accounts predict that performance-contingent rewards are detrimental to intrinsic motivation. Our finding that rewards specifically tied to level of performance (surpassing a score, exceeding others) do not undermine measures of intrinsic motivation is incompatible with the claims of these theories. On the other hand, cognitive evaluation theory could handle the pattern of results if rewards offered for doing a task or for doing well were shown to be controlling and rewards tied to performance level were shown to enhance perceptions of competence. In this case, rewards tied to performance level would also be controlling, but the competence information based on the reward procedure would override the controlling aspect of reward. Although this is a possible way to map cognitive evaluation theory onto the current results, an analysis by Eisenberger, Pierce, and Cameron (1999) indicates that reward contingencies enhance perceptions of autonomy or self-determination, a finding that is in direct contrast to predictions of cognitive evaluation theory. Thus, cognitive evaluation theory would require modification in order to handle positive effects of rewards tied to level of performance and the fact that reward contingencies can increase perceptions of self-determination.

Practical Implications

The findings from our reanalysis are in accord with a retrospective survey on the effects of extrinsic reward offered to children for reading. Flora and Flora (1999) examined the effects of parental pay for reading as well as participation in the ‘‘Book It’’ reading program sponsored by Pizza Hut. The ‘‘Book It’’ program involved over 22 million children in Australia, Canada, and the United States. The children set reading goals and were rewarded with coupons redeemable for pizzas if they met their objectives. Flora and Flora's findings indicate that neither offers of money or pizzas negatively affected reading or intrinsic motivation for reading in everyday life. These results indicate that the findings from our meta-analysis have external validity. That is, in both laboratory situations and in everyday settings, rewards offered contingent on meeting a specific level of performance do not negatively affect intrinsic motivation.

Our analysis shows that rewards can be used effectively to enhance interest without disrupting performance of an activity in a free-choice setting. These findings are given more importance in light of the fact that the group-design experiments on rewards and intrinsic motivation were primarily designed to detect detrimental effects. The reward contingencies examined in this literature can be viewed as a subset of the many possible arrangements of the use of reward in everyday life. Rewards can be arranged to shape performance progressively (Schunk, 1983, 1984), to establish interest in activities that lack initial interest (Bandura, 1986), and to maintain or enhance effort and persistence at a task (Eisenberger, 1992). Further research is necessary to show when and under what conditions rewards have positive effects on human behavior. What is clear at this time is that rewards do not inevitably have pervasive negative effects on intrinsic motivation. Nonetheless, the myth continues.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analyses.

*Adorney, K. M. (1983). *Facilitating and undermining intrinsic motivation: A test of attribution, cognitive evaluation and competence theories*. Unpublished doctoral dissertation, Columbia University.

*Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology, 50,* 14–23.

*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.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Baron, A., & Derenne, A. (2000). Quantitative summaries of single-subject studies: What do group comparisons tell us about individual performances? *The Behavior Analyst, 23,* 101–106.

*Bartelme, L.A. (1983). *The effects of choice and rewards on intrinsic motivation and performance.* Unpublished doctoral dissertation, University of Iowa.

Bates, J. A. (1979). Extrinsic reward and intrinsic motivation: A review with implications for the classroom. *Review of Educational Research, 49,* 557–576.

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 reinforcement of intrinsic motivation for sex-linked tasks. *Sex Roles, 10,* 369–386.

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., & Barrett, M. (1985). Performance and motivational deficits of helplessness: The role of motivational orientations. *Journal of Personality and Social Psychology, 49,* 1753–1761.

*Boggiano, A. K., Harackiewicz, J. M., Besette, 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,* 49–61.

*Boggiano, A. K., Main, D. S., & Katz, P. A. (1988). Children's preference for challenge: The role of perceived competence and control. *Journal of Personality and Social Psychology, 54,* 131–141.

*Boggiano, A. K., & Ruble, D. N. (1979). Competence and the overjustification effect: A developmental study. *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.

*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. *The Behavior Analyst, 19,* 237–255.

*Carton, J. S., & Nowicki, S. (1998). Should behavior therapists stop using reinforcement? A reexamination of the undermining effect of reinforcement on intrinsic motivation. *Behavior Therapy, 29,* 65–86.

*Chung, K. T. (1995). The effects of extrinsic reinforcement on intrinsic motivation amongst mildly mentally handicapped children. *Curriculum Forum, 4,* 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 doctoral dissertation, Ohio State University.

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., & Lonkey, 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.* New York: Academic Press.

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

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

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

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

*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., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125,* 627–668.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

*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.

Derenne, A., & Baron, A. (1999). Human sensitivity to reinforcement: A comment on Kollins, Newland, and Critchfield's (1997) quantitative literature review. *The Behavior Analyst, 22,* 35–41.

Dickinson, A. M. (1989). The detrimental effects of extrinsic reinforcement on ''intrinsic motivation.'' *The 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.

*Earn, B. M. (1982). Intrinsic motivation as a function of extrinsic financial rewards and subjects' locus of control. *Journal of Personality, 50,* 360–373.

*Effron, 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. (1992). Learned industriousness. *Psychological Review, 99,* 248–267.

*Eisenberger, R. (1999). *Effects of gender and* *cooperative-versus-individual performance contingent reward on intrinsic motivation*. Unpublished manuscript.

Eisenberger, R., & Cameron, J. (1996). The detrimental effects of reward: Myth or reality? *American Psychologist, 51,* 1153–1166.

Eisenberger, R., Pierce, W. D., & Cameron, J. (1999). Effects of reward on intrinsic motivation: Negative, neutral, and positive. *Psychological Bulletin, 125,* 677–691.

*Eisenberger, R., Rhoades, L., & Cameron, J. (1999). Does pay for performance increase or decrease self-determination and intrinsic motivation? *Journal of Personality and Social Psychology, 77,* 1026–1040.

*Eisenstein, N. (1985). Effects of contractual, endogenous, or unexpected rewards on high and low interest presechoolers. *The Psychological Record, 35,* 29–39.

*Enzle, M. E., Roogeveen, 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.

*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.

Flora, S. R. (1990). Undermining intrinsic interest from the standpoint of a behaviorist. *The Psychological Record, 40,* 323–346.

Flora, S. R., & Flora, D. B. (1999). Effects of extrinsic reinforcement for reading during childhood on reported reading habits of college students. *The Psychological Record, 49,* 3–14.

*Freedman, S. M., & Phillips, J. S. (1985). The effects of situational performance constraints on intrinsic motivation and satisfaciton: 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 in intrinsic motivation. *Bulletin of the Psychonomic Society, 22,* 313–316.

*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., & 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. M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest. *Journal of Personality and Social Psychology, 47,* 287–300.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 14–50). Baltimore: The John Hopkins University Press.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hennessey, B. A., & Amabile, T. M. (1998). Reward, intrinsic motivation, and creativity. *American Psychologist, 53,* 674–675.

*Hitt, D. D., Marriott, R. G., & Esser, J. K. (1992). Effects of delayed rewards and task interest on intrinsic motivation. *Basic and Applied Social Psychology, 13,* 405–414.

*Hom, H. L. (1987). A methodological note: Time of participation effects on intrinsic motivation. *Personality and Social Psychology Bulletin, 13,* 210–215.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage Publications.

*Hyman, C. (1985). *Reward contingency, standards, and intrinsic motivation.* Unpublished doctoral dissertation, City University of New York.

*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., 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.

Kollins, S. H., Newland, M. C., & Critchfield, T. S. (1997). Human sensitivity to reinforcement in operant choice: How much do consequences matter? *Psychonomic Bulletin and Review, 4,* 208–220. Erratum: *Psychonomic Bulletin and Review, 4,* 431.

*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. S., Shabatai, L., & Zaksh, D. (1975). Can money enhance intrinsic motivation?: A test of the content-consequence 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. (1998). A whole much less than the sum of its parts. *American Psychologist, 53,* 675–676.

*Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: 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.

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

*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., & 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.

Morgan, M. (1984). Reward-induced decrements and increments in intrinsic motivation. *Review of Educational Research, 54,* 5–30.

*Mynatt, C., Oakley, 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.

*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 re-

ward 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., & Svartdal, F. (1996). Effect of reward on subjective autonomy and interest when initial interest is low. *Psychological Record, 46,* 319–331.
*Pallak, S. R., Costomotis, 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.
Porter, L. W., & Lawler, E. E. (1968). *Managerial attitudes and performance.* Homewood, IL: Irwin-Dorsey.
*Pretty, G. H., & Seligman, C. (1984). Affect and the overjustification effect. *Journal of Personality and Social Psychology, 46,* 1241–1253.
*Reiss, S., & Sushinsky, L. W. (1975). Overjustification, competing responses and the acquisition of intrinsic interest. *Journal of Personality and Social Psychology, 31,* 1116–1125.
*Rosenfield, D., Folger, R., & Adelman, H. F. (1980). When rewards reflect competence: A qualification of the overjustification effect. *Journal of Personality and Social Psychology, 39,* 368–376.
*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, 33,* 442–447.
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., Mims, B., & 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.
*Salincik, 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., & Harackiewicz, J.M. (1998). ''Reality'' is complicated. *American Psychologist, 53,* 673–674.
*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.
Schunk, D. H. (1983). Reward contingencies and the development of children's skills and self-efficacy. *Journal of Educational Psychology, 75,* 511–518.
Schunk, D. H. (1984). Enhancing self-efficacy and achievement through rewards and goals: Motivational and informational effects. *Journal of Educational Research, 78,* 29–34.
Schwarzer, R. (1991). Meta: Programs for secondary data analysis. MS-DOS Version 5.0 [computer program]. Dubuque, IA: Wm. C. Brown.
*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-Kauffman, S. E. (1990). *The effects of reward contingency and type of learning experience on intrinsic motivation.* Unpublished doctoral dissertation, City University of New York.

Skaggs, K. J., Dickinson, A. M., & O'Connor, K. A. (1992). The use of concurrent schedules to evaluate the effects of extrinsic rewards on intrinsic motivation: A replication. *Journal of Organizational Behavior Management, 12,* 45–83.

*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,* 565–573.

*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., Jr., & Pittman, T. S. (1977). Moderating influence of verbal cues on intrinsic motivation. *Child Development, 48,* 1128–1132.

*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, D. (1993). Need for cognition and desire for control as moderators of extrinsic reward effects: A person × situation approach to the study of intrinsic motivation. *Journal of Personality and Social Psychology, 64,* 987–999.

*Tripathi, K. N. (1991). Effect of contingency and timing of reward on intrinsic motivation. *Journal of General Psychology, 118,* 97–105.

*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,* 241–246.

*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., & 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.

Vroom, V. (1964). *Work and motivation.* New York: Wiley.

*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.

*Wicker, F. W., Brown, G., Wiehe, J. A., & Shim, W. Y. (1990). Moods, goals, and measures of intrinsic motivation. *Journal of Psychology, 124,* 75–86.

*Williams, B. W. (1980). Reinforcement, behavior constraint and the overjustification effect. *Journal of Personality and Social Psychology, 39,* 599–614.

*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.

APPENDIX A

Following is a list of studies in which there were differences between our treatment of the study and Deci et al.'s (1999). If the difference was in terms of effect size and our effect sizes differed by more than 0.10 in either direction from Deci et al.'s, an explanation of how we calculated the effect size is given

Study	Differences
Adorney (1983)[a]	Not included by Deci et al. The study assessed the effects of tangible rewards offered for surpassing a score.
R. Anderson, Manoogian, and Reznick (1976)	For tangible reward, Deci et al. noted in Appendix A of their article that there was only one appropriate control group comparison to sue (we used that group). However, with verbal reward, Deci et al. used the inappropriate control groups for their comparison.
S. Anderson and Rodin (1989)	For verbal reward, Deci et al. did not include the free-choice measure (reported on p. 461 of the original article).
Bartelme (1983)[a]	Deci et al. did not include the free-choice measure in their analysis of performance-contingent rewards.
Boggiano and Barrett (1985)	Not included by Deci et al. The study assessed the effects of verbal reward on intrinsic motivation.
Boggiano, Main, and Katz (1988)	Not included by Deci et al. The study assessed the effects of verbal reward on intrinsic motivation.
Boggiano and Ruble (1979)	Our free-choice effect size for tangible expected rewards offered for doing (−0.61) was calculated from means and *SD*s reported in the article; it is not clear how Deci et al. obtained their estimate (−0.94).
Brennan and Glover (1980)	Our free-choice effect size for tangible rewards offered for doing (1.0) was calculated from the *F* value reported in the article; it is not clear how Deci et al. obtained their estimate (0.52).
Brewer (1980)[a]	Our free-choice effect size (−0.08) for rewards offered for doing well (Deci et al. label as performance contingent) was calculated from means and the mean square error reported in the dissertation; it is not clear how Deci et al. obtained their estimate (−0.20).
Calder and Staw (1975)	For low-interest tasks on the self-report measure (0.61) and for tangible rewards offered for finishing task (−0.46), we calculated the effect sizes from the *F* value reported in the article; it is not clear how Deci et al. obtained their estimates (1.01 for low-interest task and −0.76 for expected tangible reward).
Carton and Nowicki (1998) Experiments 1 and 2	Recent studies not included by Deci et al. The studies assessed the effects of tangible rewards offered for each unit solved.
D. S. Cohen (1974)[a]	The study was a 2 × 2 × 2 design; 2 levels of task choice (choice, no choice), 2 levels of monetary reward (money, no money) and 2 levels of verbal praise (praise, no praise). For verbal reward, our effect size is based on the main effect of verbal praise; Deci et al. omitted conditions, and their effect size is based on verbal praise versus no praise for the no-money no-choice condition only.
Crino and White (1982)	For verbal reward, Deci et al. reported a free-choice effect size; there was no free-choice measure. This study also included a low-interest task; the effect size for this was not included in Deci et al.'s analysis of low-interest tasks.

APPENDIX A

Continued

Study	Differences
Dimitroff (1984)[a]	For tangible expected rewards for doing, our effect sizes are based on the whole sample; Deci et al.'s effect sizes are based on part of the sample. In addition, Dimitroff reports no difference between the rewarded and control groups (identical means) for the self-report measure; Deci et al. report −0.26.
Eisenberger (1999)	Recent study not included by Deci et al.; included in Eisenberger, Pierce, and Cameron (1999). The study assessed the effects of tangible rewards offered for surpassing a score.
Eisenberger, Rhoades, and Cameron (1999)	Recent study not included by Deci et al. The study assessed the effects of tangible rewards offered for surpassing a score and tangible rewards offered for exceeding others.
Eisenstein (1985)	This study included a low-interest task; Deci et al. did not include the effect size for free-choice unexpected reward in their analysis of low-interest tasks.
Feehan and Enzle (1991) Experiment 1	Mislabeled by Deci et al. as Experiment 2.
Feehan and Enzle (1991) Experiment 2	Not included by Deci et al.
Freedman and Phillips (1985)	Not included by Deci et al. who stated that the tasks used were uninteresting. Close inspection of the study, however, suggests that both a high- and low-interest task were used (see comments on Phillips & Freedman, 1985). Deci et al. did not include the study in their primary analysis of the effects of reward on high-interest tasks or in their supplemental analysis of reward effects on low-interest tasks.
G. S. Goldstein (1980)[a]	Deci et al. did not include the free-choice effect size.
L. W. Goldstein (1977)[a]	The study was a 2 × 2 design; 2 levels of tangible reward (reward, no reward) and 2 levels of verbal feedback (feedback, no feedback). For verbal reward, our effect size is based on the main effect of verbal feedback; Deci et al. compared feedback to no feedback in the no-tangible-reward condition only.
Griffith (1984)[a]	In this study, there were two rewarded groups. In one group, participation was individual, in the other, it was group participation. In their analysis of low-interest tasks and in their analysis of engagement-contingent reward, Deci et al. included only the effects of reward for the individual context (but their sample size indicates they may have actually used both). We used both contexts and calculated effect sizes from *F* values, means, and *SD*s.
Harackiewicz et al. (1984) Experiments 1 and 3	Our effect sizes for rewards offered for exceeding others are based on a comparison of rewarded groups to a control group given feedback and to a control group given feedback with performance objectives comparable to the reward group. Deci et al. omitted one of the control groups and based their effect size on a comparison to the feedback control only.
Hom (1987) Experiments 1 and 2	Excluded by Deci et al. who stated that there was too little information to include. We included Experiment 2 in the verbal reward category (there was sufficient information); Experiment 1 was included in the overall reward analysis and in the analysis of tangible rewards. There was no information about reward expectancy or contingency; we did not include the study in these analyses.

APPENDIX A

Continued

Study	Differences
Kast and Connor (1988)	The study was concerned with the effects of verbal feedback on intrinsic motivation. There were 180 participants in the rewarded group and 60 in the control group; Deci et al. report 90 and 30.
Kruglanski et al. (1975) Experiment 2	Deci et al. excluded an analysis of the effects of rewards on the self-report measure for one of the tasks (stock market game).
Lepper, Sagotsky, Dafoe, and Greene (1982) Experiment 3	Our free-choice effect size for rewards offered for doing (−0.13) was calculated from means and *SD*s; it is not clear how Deci et al. obtained their estimate (−0.50).
McLoyd (1979)	Deci et al. did not include the self-report measure in their analysis of completion-contingent rewards.
Mynatt, Oakley, Piccione, Margolis, and Arkkelin (1978)	Our free-choice effect size for expected tangible rewards offered for doing (0.19) is based on between-group differences; Deci et al. used a within-group comparison (−0.11).
Okano (1981) Experiment 2	For task-noncontingent reward, our free-choice effect size (−0.47) was calculated from means and *SD*s reported in the article; it is not clear how Deci et al. obtained their estimate (−0.84). On the self-report measure, Deci et al.'s effect size should be negative, not positive.
Overskeid and Svartdal (1996) Experiments 1 and 2	Excluded by Deci et al. because the task was of low initial interest. This study should have been included in Deci et al.'s analysis of reward effects on low-interest tasks.
Patrick (1985)[a]	There is not enough information in the study to calculate a free-choice or self-report effect size.
Phillips and Freedman (1985)	Excluded by Deci et al. who stated that the tasks used were uninteresting. Two tasks were used in this study; one was rated above the median on a 7-point scale (3.8), the other was rated below the median (2.7) (see p. 307 of the original article). Freedman and Phillips (1985) was also in accord with this scale. This study was not included in either Deci et al.'s primary analysis of the effects of reward on high-interest tasks or their supplemental analysis of reward effects on low-interest tasks.
Picek (1976)	In this study, there were three reward conditions; one group received a reward for doing the task (we included this in our analysis of rewards offered for doing). In the other two reward groups (one was labeled performance contingent by Deci et al.), only half the participants in the rewarded conditions were actually offered and given a reward. Because not all participants were offered a reward, we excluded these two conditions from our analysis.
Pittman, Cooper, and Smith (1977)	Our free-choice effect size (−0.50) for rewards offered for surpassing a score (Deci et al. label the study performance contingent) was calculated from the *p* value reported in the article; it is not clear how Deci et al. obtained their estimate (−1.46).
Pittman, Emery, and Boggiano (1982) Experiment 1	Deci et al. did not include the self-report measure for engagement-contingent rewards; a self-report measure is reported in the article on page 792.
Pretty and Seligman (1984) Experiment 1	The study was 3 × 3 design; 3 levels of tangible reward (expected, unexpected, no reward) and 3 levels of feedback (positive, negative, no feedback). For verbal reward, we compared positive feedback to no feedback across all reward conditions; Deci et al. compared positive feedback to no feedback in the no-reward condition only.

APPENDIX A

Continued

Study	Differences
Ross (1975) Experiment 1	For rewards offered for doing, our free-choice effect size (0.01) was calculated from *t* values reported in the article; it is not clear how Deci et al. obtained their estimate (−0.18).
Shapira (1976)	Excluded by Deci et al. who stated that rewarded participants worked on an easier task than the control group. Both groups worked on Soma puzzles. Because task difficulty was not a variable assessed in our meta-analysis, we included this study.
A. T. Smith (1980)[a]	Deci et al. did not include the free-choice effect size for verbal reward or the effect size for low-interest tasks in their analysis of low-interest tasks.
T. W. Smith and Pittman (1978)	Our free-choice effect size (−0.56) for rewards offered for surpassing a score (Deci et al. label the study performance contingent) was calculated from the *p* value reported in the article; it is not clear how Deci et al. obtained their estimate (−94).
W. E. Smith (1975)[a]	For verbal reward, Deci et al. used only part of the sample. There were two verbal reward groups ($n = 40$) and a control condition ($n = 20$); Deci et al. report 20, 20. For unexpected reward, Deci et al. also used only part of the sample.
Thompson, Chaiken, and Hazlewood (1993)	For engagement-contingent rewards (or rewards offered for doing a task), Deci et al. report the self-report effect size as the free-choice effect size and did not give an effect size for the self-report measure.
Tripathi (1991)	Not included by Deci et al. The study assessed the effects of tangible rewards offered for doing a task and for surpassing a score.
Tripathi and Agarwal (1988)	For rewards offered for doing, our free-choice effect size (0.34) was calculated from *F* values reported in the article; it is not clear how Deci et al. obtained their estimate (0.00).
Weiner (1980)	Our free-choice effect size (0.35) for rewards offered per unit solved (Deci et al. labeled the study completion contingent) was calculated from means and *SD*s reported in the article; it is not clear how Deci et al. obtained their estimate (0.20).
Wicker et al. (1990)	Excluded by Deci et al. who reported that the article was not about the effects of rewards on intrinsic motivation and none of the appropriate statistics were available. We wrote to the first author and obtained the data for the free-choice and self-report measures.
Williams (1980)	For rewards offered for doing, our free-choice effect size is based on the whole sample. Although Deci et al. report the entire sample, their effect size appears to be based on only part of the sample.
Wilson (1978)[a]	In this study, there were two reward groups (one received \$2.50; the other group received \$0.50); Deci et al. included only one of the reward groups in their analysis of low-interest tasks and in their analysis of engagement-contingent rewards. We included both. For other studies that used two different magnitudes of rewards, Deci et al. included both (e.g., McLoyd, 1979).

[a] Unpublished doctoral dissertation.

APPENDIX B

A comparison of our classification of reward contingencies with Deci, Koestner, and Ryan (1999)

Our classification of the reward contingency	Differences between our categories and Deci et al.'s (1999)
Task-noncontingent reward	Also labeled task noncontingent by Deci et al. The same set of studies was used in both analyses.
Rewards offered for doing well	All studies included in our analysis of this category were labeled performance contingent by Deci et al.
Rewards offered for doing the task	Studies included in our analysis of this category were labeled engagement contingent by Deci et al. We included L. W. Goldstein (1977) in this category. In this study participants were offered a reward to take pictures; no instructions were given about completing the task and there was no requirement to do well or to achieve any specific standard. Deci et al. labeled this study as completion contingent.
Rewards offered for finishing or completing a task	Studies in our analysis of this category were labeled completion contingent by Deci et al. We included Tripathi and Agarwal (1985) in this category. In this study rewarded participants were told they could earn a reward if they completed the task even if all solutions were not correct. Deci et al. labeled this study as engagement contingent.
Rewards offered for each puzzle or problem solved	Most of the studies in our analysis of this category were labeled completion contingent by Deci et al. Five studies in this category were labeled performance contingent by Deci et al. (Bartelme, 1983; D. S. Cohen, 1974; Effron, 1976; Lee, 1982; Weiner & Mander, 1978). In each of these studies, participants were offered pay for each point earned on a puzzle-solving task, each word found in a word-search task, each code completed on a decoding task, or each correct answer on a matching-to-sample task.
Rewards offered for surpassing a score	All studies in our analysis of this category were labeled performance contingent by Deci et al.
Rewards offered for exceeding a norm	All studies in our analysis of this category were labeled performance contingent by Deci et al.

APPENDIX C

Studies included in the analysis of the effects of rewards on intrinsic motivation for tasks with low initial interest

Study	Reward type	Reward expectancy	Reward contingency	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)
Calder and Staw (1975)	T	E	For finishing task	10	10		0.61
Chung (1995)	T	E	For doing task	5	5	1.93	
	T	E	Insufficient information	5	5	1.22	
Crino and White (1982)	V	U	Per unit solved	10	5		−0.05
	V	U	Yoked per unit	10	5		0.32
Daniel and Esser (1980)	T	E	For doing quickly	16	16	−0.28	0.08
Eisenstein (1985)	T	U	For finishing task	6	6	0.62	
	T	E	For finishing	16	6	0.22	
Freedman and Phillips (1985)	T	E	Per unit solved	24	25		0.24
	T	E	For finishing task	26	25		0.53
Griffith (1984)[a]	T	E	For doing task	44	44	0.25	
Hamner and Foster (1975)	T	E	For doing task	16	15		−0.28
	T	E	Per unit solved	19	15		0.52
Hitt, Marriott, and Esser (1992)	T	E	For doing task	30	15	0.57	−0.16
Loveland and Olley (1979)	T	E	For doing task	6	6	1.20	
McLoyd (1979)	T	E	For finishing task	18	9	0.61	0.00
Mynatt et al. (1978)	T	E	For doing task	5	5	1.35	
Newman and Layton (1984)	T	E	For doing task	20	10	0.41	
Overskeid and Svartdal (1996) Experiment 1	T	E	For doing task	10	10	−0.29	0.39
Overskeid and Svartdal (1996) Experiment 2	T	E	For doing task	64	32		−0.15
Phillips and Freedman (1985)	T	E	For finishing task	12	12		0.63
	T	E	Per unit solved	12	12		−0.10
A. T. Smith (1980)[a]	T	E	For doing task	21	27	0.04	
	V	U	For doing task	22	26	0.17	
Wilson (1978)[a]	T	E	For doing task	46	23	−0.03	0.12

Note. T = tangible reward, V = verbal reward, E = expected, U = unexpected, N_E = sample size of experimental group, N_C = sample size of control group.
[a] Unpublished doctoral dissertation.

APPENDIX D

Studies included in the analysis of the effects of verbal rewards on intrinsic motivation for tasks with high initial interest

Study	Reward expectancy	Reward contingency	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)
R. Anderson et al. (1976)	U	For doing task	18	19	0.40	
S. Anderson and Rodin (1989)	U	For doing task	10	10	0.20	0.40
Blanck, Reis, and Jackson (1984) Experiment 1	U	For doing task	70	69	0.56	0.69
Blanck et al. (1984) Experiment 2	U	For doing task	12	12	0.73	0.00
Boggiano and Barrett (1985)	U	For doing task	18	18	0.35	
Boggiano et al. (1988)	U	For doing task	66	34	0.42	
Butler (1987)	E	For doing task	50	50		1.59[b]
D. S. Cohen (1974)[a]	U	For doing task	52	52	0.07	0.42
Crino and White (1982)	U	Per unit solved	10	5		0.05
	U	Yoked per unit	10	5		−0.79
Danner and Lonkey (1981)	U	For doing task	30	30	−0.10	−0.08
Deci (1971) Experiment 3	U	For doing task	12	12	0.82	0.00
Deci (1972b)	U	For doing task	48	48	0.29	
Deci, Cascio, and Krusell (1975)	No information	No information	32	32	0.02	
Dollinger and Thelen (1978)	E	For doing well	12	12	−0.07	0.00
Effron (1976)[a]	U	For doing task	15	13		0.89
L. W. Goldstein (1977)[a]	U	For doing task	32	32	0.77	0.12
Harackiewicz (1979)	U	For doing task	31	31		0.59
Hom (1987) Experiment 2	No information	No information	28	28	−0.37	
Kast and Connor (1988)	U	For doing task	180	60		−0.46[b]
Koestner, Zuckerman, and Koestner (1987)	U	For doing task	35	18	0.51	0.00
Orlick and Mosher (1978)	U	For doing task	11	12	−0.34	
Pallak, Costomotis, Sroka,	U	For doing task	16	12	−0.47	
and Pittman (1982)	E	For doing task	14	12	0.32	
Pittman, Davey, Alafat, Wetherill, and Kramer (1980)	U	For doing task	24	12	0.80	
Pretty and Seligman (1984) Experiment 1	U	For doing task	30	30	0.35	0.46
Ryan, Mims, and Koestner (1983)	E	For doing task	32	16	0.53	0.00
Sansone (1986)	U	For doing task	44	11		0.68
Sansone (1989)	E	For doing task	82	41		0.46
Sansone, Sachau, and Weir (1989)	U	For doing task	40	40		0.12
Shanab, Peterson, Dargahi, and Deroian (1981)	U	For doing task	20	20	0.64	0.43
A. T. Smith (1980)[a]	U	For doing task	21	27	0.24	
W. E. Smith (1975)[a]	U	For learning	20	20	0.04	0.00
Tripathi and Agarwal (1985)	E	For doing task	20	20	1.61	0.48
Vallerand (1983)	E	For doing task	40	10		1.98[b]
Vallerand and Reid (1984)	E	For doing task	28	28		0.53
Zinser, Young, and King (1982)	U	For doing task	64	32	0.08	

Note. U = unexpected, E = expected, N_E = sample size of experimental group, N_C = sample size of control group.
[a] Unpublished doctoral dissertation.
[b] Outliers in the data set.

APPENDIX E

Studies included in the analysis of unexpected tangible rewards on intrinsic motivation for tasks with high initial interest for all reward contingencies

Study	Reward contingency	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)
Eisenstein (1985)	For finishing	10	10	0.46	
Greene and Lepper (1974)	For doing well	26	15	0.14	
Harackiewicz et al. (1984) Experiment 2	Exceeding others	15	15	0.44[b]	0.15
Kruglanski, Alon, and Lewis (1972)	For winning	36	33		−0.65
Lepper et al. (1973)	For doing task	18	15	0.12	
Orlick and Mosher (1978)	For doing well	12	12	−1.28	
Pallak et al. (1982)	For doing task	15	12	−0.43	
Pretty and Seligman (1984) Experiment 1	For doing task	30	30	0.06	0.42
Pretty and Seligman (1984) Experiment 2	For doing task	30	30	0.06	0.38
W. E. Smith (1975)[a]	For learning	40	40	0.06	0.00

Note. N_E = sample size of experimental group, N_C = sample size of control group.

[a] Unpublished doctoral dissertation.

[b] Effect sizes based on performance measures on the task during the free-choice period (e.g., number of balls played in a pinball game, number of trials initiated in a labyrinth game, number of words found in a word search game). See text for details.

APPENDIX F

Studies included in the analysis of expected tangible rewards on intrinsic motivation for tasks with high initial interest, listed by reward contingency

Study	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)	Reward delivery
Task noncontingent					
Dafoe (1985)[a]	25	28	−0.20	0.73	
Deci (1972a)	24	16	0.08		
Earn (1982)	40	20	−0.28	0.18	
Kruglanski, Friedman, and Zeevi (1971)	16	16		−0.69	
Okano (1981) Experiment 2	11	11	−0.47	−0.27	
Pittman et al. (1982) Experiment 1	10	10	0.26	0.00	
Ross, Karniol, and Rothstein (1976)	12	12	0.44		
Swann and Pittman (1977) Experiment 1	20	20	−0.21		
Wimperis and Farr (1979)	16	16		0.56	
Rewards offered for doing task					
Amabile, Hennessey, and Grossman (1986) Experiment 1	56	57	0.00	0.00	
Amabile et al. (1986) Experiment 3	30	30		0.00	
R. Anderson et al. (1976)	36	19	−0.53		
Arnold (1976)	17	36		0.00	
Arnold (1985)	13	16		−0.04	
Boggiano and Ruble (1979)	20	20	−0.61		
Boggiano, Havackiewicz, Besette, and Main (1985)	26	13	−0.79		
Boggiano, Ruble, and Pittman (1982)	81	84	0.28		
Brennan and Glover (1980)	20	19	1.00[b]		
Brewer (1980)[a]	24	24	−0.13	0.12	
Chung (1995)	5	5	−1.61[b]		
Danner and Lonkey (1981)	30	30	−1.33[b]	−1.23	
DeLoach, Griffith, and LaBarba (1983)	26	26	0.00		
Dimitroff (1984)[a]	108	36	−0.27	0.00	
Effron (1976)[a]	12	13		0.19	
Fabes, McCullers, and Hom (1986)	24	24	0.06	−0.14	
Fabes et al. (1988)	14	14	−1.34[b]	−0.76	
Fabes, Fultz, Eisenberg, May-Plumlee, and Christopher (1989)	15	14	−0.73		
Feehan and Enzle (1991) Experiment 1	24	12	−0.97		
L. W. Goldstein (1977)[a]	16	16	−0.99	−0.87	
Greene and Lepper (1974)	15	15	−0.70		
Griffith (1984)[a]	44	44	−0.23		
Hamner and Foster (1975)	15	15		−0.14	
Harackiewicz (1979)	31	31		−0.38	
Hitt et al. (1992)	30	15	−0.82	−0.47	
Hyman (1985)[a]	32	32	−0.42		
Karniol and Ross (1977)	17	20	−0.08		
Lepper et al. (1973)	18	15	−0.72		
Lepper et al. (1982)	32	32	−0.13		
Loveland and Olley (1979)	6	6	−1.20		
Morgan (1981) Experiment 1	27	27	−0.98	−0.31	
Morgan (1981) Experiment 2	20	20	−0.77	0.04	
Morgan (1983) Experiment 1	40	40	−1.94[b]		
	40	20		−0.54	

APPENDIX F

Continued

Study	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)	Reward delivery
Morgan (1983) Experiment 2	20	20	−0.66	0.00	
Mynatt et al. (1978)	5	5	0.19		
Newman and Layton (1984)	20	10	−0.37		
Ogilvie and Prior (1982)	26	26	−0.08		
Okano (1981) Experiment 1	15	15	−0.99	−0.45	
Okano (1981) Experiment 2	10	11	−1.31[b]	0.00	
Patrick (1985)[a]	33	31	0.00	0.00	
Perry, Bussey, and Redman (1977)	32	32	−0.43	−0.21	
Picek (1976)[a]	10	10	0.00	−0.65	
Pittman et al. (1982) Experiment 1	10	10	0.17	0.00	
Pittman et al. (1982) Experiment 2	27	27	−0.05		
Pretty and Seligman (1984) Experiment 1	30	30	−0.75	−0.05	
Pretty and Seligman (1984) Experiment 2	30	30	−0.13	−0.16	
Reiss and Sushinsky (1975)	16	16	−0.83		
Ross (1975) Experiment 1	40	20	0.01	−0.45	
Ross (1975) Experiment 2	52	14	−0.66	0.00	
Ross et al. (1976)	12	12	−0.64		
Ryan et al. (1983)	16	16	−0.35	0.00	
Sarafino (1984)	85	15	−0.41	0.00	
Shiffman-Kauffman (1990)[a]	20	20	0.06	−0.04	
A. T. Smith (1980)[a]	21	27	−0.82		
Swann and Pittman (1977) Experiment 1	20	20	−0.78		
Swann and Pittman (1977) Experiment 2	26	13	−1.01		
Thompson et al. (1993)	34	33	−0.003	0.14	
Tripathi (1991)	20	5	0.00	0.00	
Tripathi and Agarwal (1988)	20	10	0.34[b]	0.72	
Weiner and Mander (1978)	30	30	−0.34	0.00	
Williams (1980)	24	24	0.18	0.00	
Wilson (1978)[a]	46	23	−0.06	−0.01	
Yuen (1984)[a]	60	60	−0.40	−0.12	
Rewards offered for ‘‘doing well’’ or ‘‘doing a good job’’ on the task					
Brewer (1980)[a]	48	24	−0.08	0.12	M
Dafoe (1985)[a]	26	28	0.00	0.59	M
Dollinger and Thelen (1978)	36	12	−0.55	0.00	L
Enzle, Roogeveen, and Look (1991)	40	10	−0.53		M
Fabes (1987) Experiment 1	18	19	−0.87		M
L. W. Goldstein (1977)[a]	16	32	−0.08	−0.48	M
Greene and Lepper (1974)	15	15	−0.57		M
Hyman (1985)[a]	16	16	0.11		M
Orlick and Mosher (1978)	14	12	−0.53		M
Pallak et al. (1982)	15	12	−0.17		M
Ryan et al. (1983)	32	32	−0.46	0.00	M
Taub and Dollinger (1975)	124	124		0.00	NI
Rewards offered for completing a task					
Calder and Staw (1975)	10	10		−0.46	M
Eisenstein (1985)	18	10	−0.53		M
Fabes (1987) Experiment 1	19	19	−0.82		M
Fabes (1987) Experiment 2	14	14	−0.45		M
Freedman and Phillips (1985)	26	22		0.94	M
Griffith, DeLoach, and LaBarba (1984)	64	32	0.00		M
McLoyd (1979)	18	9	−1.04	0.00	M
Phillips and Freedman (1985)	12	12		0.74	M

APPENDIX F

Continued

Study	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)	Reward delivery
Staw, Calder, Hess, and Sanderlands (1980)	47	46		0.19	M
Tripathi and Agarwal (1985)	20	20	0.41	0.54	M
Rewards offered for each problem, puzzle, or unit solved					
Arkes (1979)	32	32	−0.16	0.03	M
Arnold (1985)	13	16		−0.05	L
Bartelme (1983)[a]	35	34	0.04[c]	0.03	M
Boggiano et al. (1985)	26	13	−0.10		M
Brockner and Vasta (1981)	26	26	−0.37	−0.58	L
Carton and Nowicki (1998) Experiment 1	44	22	0.36[c]		L
Carton and Nowicki (1998) Experiment 2	40	20	0.20[c]	0.71	L
D. S. Cohen (1974)[a]	52	52	−0.18	0.13	L
Deci (1971) Experiment 1	12	12	−0.54	0.00	L
Deci (1972b)	64	32	0.33		L
Effron (1976)[a]	43	28		−0.04	L
Feehan and Enzle (1991) Experiment 2	30	15	0.31[c]		M
Freedman and Phillips (1985)	23	22		1.12	L
G. S. Goldstein (1980)[a]	14	14	−0.32	0.68	L
Hamner and Foster (1975)	18	15		−0.21	L
Kruglanski et al. (1975) Experiment 1	24	24		1.15[b]	M
Lee (1982)[a]	40	40	−0.36[c]	0.35	M
Liberty (1986) Experiment 1[a]	23	23	−0.86[c]	−0.34	L
Liberty (1986) Experiment 2[a]	44	42	−0.22[c]	0.04	L
McGraw and McCullers (1979)	20	20		−0.04	NI
Phillips and Freedman (1985)	12	12		0.77	L
Porac and Meindl (1982)	20	20	−0.78		L
Shapira (1976)	30	30		0.41	L
Sorensen and Maehr (1976)	20	20	−0.54		L
Vasta and Stirpe (1979)	4	5	−0.16		L
Weiner (1980)	24	24	0.35	0.00	M
Weiner and Mander (1978)	30	30	−0.54	0.00	L
Wicker et al. (1990)	29	29	−0.46	0.18	L
Wimperis and Farr (1979)	16	16		1.36[b]	NI
Rewards offered for meeting a specific standard or surpassing a score					
Adorney (1983)[a]	35	36	0.39	0.48	L
Bartelme (1983)[a]	35	34	0.19[c]	−0.03	M
Boggiano and Ruble (1979)	20	20	−0.17		M
Dafoe (1985)[a]	28	28	0.15	0.59	M
Eisenberger (1999)	214	316	0.08	0.31	M
Eisenberger, Rhoades, and Cameron (1999) Experiment 1	110	113	0.10	0.34	M
Harackiewicz, Abrahams, and Wageman (1987)	13	25		−0.28	M
Hyman (1985)[a]	16	16	0.04		M
Kruglanski et al. (1975) Experiment 2	40	40		0.38	M
Patrick (1985)[a]	30	31	0.00	0.00	M
Pittman et al. (1977)	60	20	−0.50[c]	−0.20	L
T. W. Smith and Pittman (1978)	66	33	−0.56[c]	0.00	L
Tripathi (1991)	20	5	0.00	0.00	M

APPENDIX F

Continued

Study	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)	Reward delivery
Rewards offered for meeting or exceeding others					
Dafoe (1985)[a]	25	28	0.00	0.59	M
Eisenberger, Rhoades, and Cameron (1999)					
Experiment 1	106	106	0.38	0.22	M
Harackiewicz (1979)	31	31		−0.87	M
Harackiewicz and Manderlink (1984)	47	47		0.33	M
Harackiewicz et al. (1984) Experiment 1	32	64	0.27	0.12	M
Harackiewicz et al. (1984) Experiment 2	15	15	−0.43[c]	−0.18	M
Harackiewicz et al. (1984) Experiment 3	26	52	0.34[c]	0.40	M
Harackiewicz et al. (1987)	11	29		0.12	M
Karniol and Ross (1977)	20	20	0.15		L/M
Luyten and Lens (1981)	10	10	−0.90	0.08	L
Rosenfield, Folger, and Adelman (1980)	30	27	0.30	0.22	L/M
Salincik (1975)	38	39	−0.34	0.01	M
Shiffman-Kauffman (1990)[a]	20	20	0.35	0.00	M
Tripathi and Agarwal (1988)	20	10	0.87	1.01	M
Weinberg and Jackson (1979)	40	40		0.00	L

Note. N_E = sample size of experimental group, N_C = sample size of control group, M = maximum reward, L = less than maximum reward, NI = not enough information.
[a] Unpublished doctoral dissertation.
[b] Outliers in the data set.
[c] Effect sizes based on performance measures during the free-choice period. See text for details.

APPENDIX G

Studies or conditions within studies included in the overall analyses of reward and tangible reward that could not be classified into reward contingencies

Study	Reward contingency	N_E	N_C	Free-choice effect size (*g*)	Self-report effect size (*g*)
Chung (1995)	Insufficient information	5	5	−1.02	
Daniel and Esser (1980)	For doing quickly	16	16	−0.75	−0.71
Hom (1987) Experiment 1	No information	26	26	0.12	0.00
Hom (1987) Experiment 2	No information	28	28	−0.37[b]	
W. E. Smith (1975)[a]	For showing learning	40	40	−0.22	0.00

Note. N_E = sample size of experimental group, N_C = sample size of control group.
[a] Unpublished doctoral dissertation.
[b] Effect size based on performance measures during the free-choice period. See text for details.