Creativity Research Journal
1999, Vol. 12, No. 4, 231–239

Effect of Expected Rewards on Children's Creativity

Mireille Joussemet and Richard Koestner
McGill University

ABSTRACT: This study examined the effect of expected rewards on children's creativity. Sixty-one female gymnasts (ages 4–17) were randomly assigned to a no-reward or expected reward condition. All participants completed both a training task that required divergent thinking (generating themes for a gymnastics gala) and a transfer task (using circles to make pictures). The reward contingency was in effect only during the training task. Creativity was assessed by (a) consensual judgment of 5 raters and (b) determining the statistical rarity of a given response for this sample of participants. Results indicated that rewards lead younger children to generate less appropriate themes on the training task and children of all ages to draw somewhat less creative pictures on the transfer task. It was also found that the consensual judgment measure of creativity was more sensitive to the age of children than was the rarity measure.

Creativity is a highly valued behavior that is perceived to be rare. Perhaps the common belief that few people have the ability to bring something new into existence is due to the fact that in the 1960s, creativity research was dominated by the trait approach (Barron, 1968). Amabile (1983) argued that "the trait approach is incomplete and that creativity is best conceptualized not as a personality trait or a general ability but as a behavior resulting from particular constellations of personal characteristics, cognitive abilities, and social environments" (p. 358). Thus, more recent research now focuses on creative products instead of people and suggests that anybody can show some creativity (Amabile, 1983). However, if this quality of work is not reserved for artists and scientists, how do the rest of us come to do original and ingenious things? Also, how can we create the conditions in which children will develop this capacity in their everyday lives?

Behaviorists and social-cognitive psychologists do not agree on the answer to these questions. Behaviorists treat creativity like any other performance dimension and state that reinforcement will increase its frequency. Social-cognitive researchers hold the view that rewarding creative behavior causes the performer to feel controlled and that this unpleasant feeling decreases creativity. Moreover, rewards can divert attention from the task, thus decreasing its creative aspect.

The social-cognitive position that reward decreases creativity has received considerable support. For example, Lepper, Greene, and Nisbett (1973) found that children who made pictures for an expected reward produced a significantly greater number of pictures than children who were not promised a reward, but that these pictures were rated as significantly lower in quality. Condry (1977) reviewed the literature and concluded that individuals given rewards "seem to work harder and produce more activity, but the activity is of lower quality, contains more errors, and is more stereotyped and less creative than the work of comparable nonrewarded participants working on the same problems" (pp. 471–472). Later reviews supported this conclusion (Deci & Ryan, 1987; Koestner & McClelland, 1990). Moreover, in their work on the social influences on creativity, Amabile and her colleagues (Amabile, Goldfarb, & Brackfield, 1990; Amabile, Hennessey, & Grossman, 1986,) identified various controlling practices besides the use of rewards that undermine creativ-

This research was funded by a grant to Richard Koestner from the Social Science and Humanities Research Council (SSHRC–Canada) and from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR–Québec).

We thank Mette Dybrad, Julie Bélanger, Annie Bérubé, Marie-Claude Besner, Jodie Engel, and Irma Mazzone for their help on this project.

Manuscript received May 15, 1998; revision received April 8, 1999; accepted May 1, 1999.

Correspondence and requests for reprints should be sent to Mireille Joussemet, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montréal, Quebec, H3A 1B1, Canada.

ity, including competition, external evaluation, and surveillance.

In a recent article, Eisenberger and Cameron (1996) adopted a behaviorist approach in reviewing the creativity literature and concluded that the detrimental effect of rewards on creativity was not well established. In fact, they argued that it was a myth. They stated that the detrimental effects are easily avoidable and that rewards can have positive effects on generalized creativity if they are properly administered. This conclusion was largely based on Eisenberger and Selbst's (1994) study in which schoolchildren either received or did not receive monetary rewards for inventing new words from strings of letters. Children's creativity was later assessed by examining the rarity of their responses on a second task that involved making pictures from circles (the transfer task). The only condition in which drawings were more creative was when children were moderately rewarded for demonstrating high divergent thinking in the first task. That is, if the children were reinforced for producing six new words per string (high divergent thinking), and if the reward was not too big or salient (so that it did not take the attention away from the task), creativity in drawings increased. This result was explained in terms of learned industriousness theory (Eisenberger, 1992), which states that "if an individual is rewarded for putting a large amount of cognitive or physical effort into a task, the sensation of high effort acquires secondary reward properties that ameliorate, to some degree, effort's innate aversiveness" (Eisenberger & Cameron, 1996, p. 1161). Reward would then increase one's readiness to put effort into a task, whether it is a creative, force, or speed task.

Eisenberger and Amabile differ in their general approach to defining, quantifying, and setting the stage for examining the impact of rewards on creativity. First, they did not define creative products in the same way. Eisenberger and his colleagues (Eisenberger & Armeli, 1997; Eisenberger & Selbst, 1994) emphasized the novelty aspect; if behavior diverges from the norm, then it is creative. By contrast, Amabile (1983) stressed the social component of creative behavior. Her operational definition relied on subjective, consensual criteria, where people agree on what is creative. These different ways to conceptualize creativity naturally lead to different ways to measure it. Indeed, Eisenberger relied on the objective measure of statistical rarity to decide if a product is creative, whereas Amabile (1982) relied on a more subjective methodology based on raters' consensual judgment. Finally, the way these two researchers "set the stage" for creativity to occur is also different. Eisenberger seemed to use constraining tasks such as school-like, paper-and-pencil tasks in which participants either generate new words from a string of letters, make pictures from printed circles, or list uses for common objects. In contrast, tasks used in Amabile's research have wider boundaries. They include more "artistic" activities in which children are asked to make drawings, write stories, or take pictures.

Research on creativity by both the behaviorists and social-cognitivists has been subject to criticism by the opposing camp. One of the criticisms of the behaviorist approach to studying creativity is that there is no way to distinguish the effects of reward as a reinforcer and the effects of the information provided by instructions and by the reward contingency itself. Thus, information could be a confounding variable. To control for the informational factors, Eisenberger and Selbst (1994) gave both the rewarded and the control group information and feedback, by saying "correct" each time the contingency was fulfilled by participants.

Regarding the social-cognitive approach, a criticism was made concerning the small number of tasks done in a study. In fact, because these researchers are concerned with the effect of expecting a reward, they usually create this expectation condition by either promising a reward or presenting a single performance–reward pairing. Behaviorists argued that when repeated presentation of the contingency is replaced by this short, abbreviated procedure, the participants may use "implicit, previously learned cues for performance that may evoke behavior different from the actual presentation of a reward contingency" (Eisenberger & Selbst, 1994, p. 1117). Some studies anticipated this criticism by using more than one task and still found decreased creativity with rewards (McGraw & McCullers, 1979; Schwartz, 1982). However, according to behaviorists, the tasks used were too simple and rewards could be obtained with a low degree of divergent thinking. Such training could have reduced participants' general tendency toward divergent thought.

This latter critique stems from the different explanations given for the decreased creativity. Unlike behaviorists, cognitively oriented psychologists do not consider reward per se as the causal variable; they consider reward expectancy to be the key element. Rewards can be given in different ways, and some are more controlling than others. If, for example, a teacher

wants to make his or her students happy and gives them a surprise, he or she can do it without necessarily making the surprise contingent on any specific behavior. If, however, a teacher has an established system of points gained for desirable behavior whereby points can be exchanged for little surprises, the reward is expected and control can operate. Children know what they have to do and for what reason.

In Eisenberger's recent experiments (Eisenberger & Armeli, 1997; Eisenberger & Selbst, 1994), it is not clear whether rewards used are expected or not by the participants. It seems that they are not, as nothing is mentioned about the contingency in the instructions. If this is the case, it is an odd method because (a) it does not represent how rewards are usually used in the real world and (b) unexpected rewards had never been seen as detrimental. Indeed, when rewards are used to increase the frequency of a behavior, usually the "rules of the games" are explained explicitly. The rewards are made available for a specific behavior fulfilling a specific demand. We understand that the contingency for creativity was not made explicitly by Eisenberger to avoid the confounding effects of information. However, if the participants do not expect any reward, then it is normal that no detrimental effect has been observed, the damaging element being absent.

Our Study

The goal of this study is to determine whether an expected reward for the display of divergent thinking on a task decreases creativity in that task and in a subsequent one. The study was designed to meet behaviorists' standards: More than one task was used, and the first one required high divergent thinking. Specifically, the training task consisted of listing many different original ideas for a show and the transfer task was identical to that used by Eisenberger and Selbst (1994): using circles to make drawings. Rewards were of medium size and were not too salient. Two measures of creativity were used: Eisenberger's measure of statistical rarity and Amabile's consensual measure.

A unique feature of our study is that we sampled a wide age range of children as participants (4–17 years old). We considered this an advantage because we wanted to examine whether our two measures of creativity, rarity of responses (following Eisenberger) and judged creativity (following Amabile) showed a positive relation to age. It would seem that a valid measure of creativity among children should reflect expected developmental patterns. Although reviews have noted that different aspects of creativity may develop at different rates (Runco, 1996), there is general agreement that older children should display greater creativity than younger ones (Torrance, 1962). For example, a large study found that the quality of creative responses improved with age (e.g., they became less stereotypical, more original, and showed more humor; Urban, 1991).

We predicted that participants in the expected reward condition would produce less creative ideas and drawings than would participants in the no-reward condition. Although Eisenberger considered the transfer task as the interesting piece of data, we planned to examine the level of creativity demonstrated in both the training and the transfer tasks.

Method

Participants

The participants were 61 female gymnasts between the ages of 4 and 17 years old. They were recruited from two gymnastics clubs in Montréal. Parents provided written consent and children were asked verbally if they wanted to participate. The girls participated in their usual training group, and these groups were randomly assigned to one of the two conditions (expected reward or no reward). Random assignment to condition was blocked by club membership and age level. Four participants were unable to generate a response to the training task and were excluded from analyses.

Materials

For the training task, numbered white sheets were provided. On the top of these sheets, space was provided so that participants could indicate their age. The only written instruction on the sheet was: "Please write down new themes for a gymnastics gala." An example (circus) was also provided. All the girls understood that each year their gymnastics club holds a gala event in which the girls participate in a performance related to some organizing theme.

For the transfer task, sheets with 12 printed circles (four rows and three columns) were provided. Each circle had a diameter of 3.8 cm. These sheets were numbered to match each participants' drawings with her themes list. No instructions were written but two drawings (a happy face and a flower) were provided as examples. This task was identical to the one used by Eisenberger and Selbst (1994).

The reward consisted of a little notebook with a matching pencil. It is a common gift for girls of this age. It is the kind of little gift a teacher or coach would give to students.

Procedure

The study was conducted in February 1998, during gymnastics class. Participants were verbally informed that the study was about creativity and would last 10 min. All instructions were given verbally in French (participants' principal language used). In the reward condition the notebooks were distributed after completing the first task. All participants remained anonymous. The sheets were numbered, and the only personal information requested was the participant's age.

Training task. The first creative task consisted of listing themes for a gymnastics gala, a relevant task for these participants, as they have to prepare this type of show each year. It was thus a real-world, open-ended divergent thinking task, as was used by Okuda, Runco, and Berger (1991). Runco and Chand (1995) stated that "real-world problems seem to be much more valuable than traditional divergent thinking tests" (p. 254) and that educators should not rely only on "assignments for which there are clearly defined solutions" (p. 249).

Each gymnast was given a sheet and a pencil to list themes or topics. An example of a possible theme was given. Each participant was asked to write down as many ideas as she could. Five minutes were allowed for this first task. The vast majority of participants (85%) produced more than 3 themes with the average number being 5.6.

In the expected reward condition, gymnasts were given the following instructions:

> I am going to give you a blank sheet. What you have to do is to find as many new topics or themes for a gymnastics gala as you can and write them down. If you write down some ideas, you will get a notebook and a pencil like this. [Experimenter shows a notebook and a pencil to the children and puts them back in a bag.] There is an example of a theme on your sheet. So, for example, a gymnastics show could talk about ... a circus. This is an example. Remember, write some new ideas for a gala. When you can't think of any more themes, just put your pencil down. Do you understand? Do you have any questions? Okay, you can start now, we'll take about five minutes to carry out this activity.

After having completed the task and having had their list "evaluated" by the researcher, gymnasts in the expected reward condition were given their notebook and pencil. They were all told, "That's good, here's your reward."

In the no-reward condition, the training task was done the same way, except that no rewards or reward information was given. Gymnasts in this condition were simply asked to list as many gala themes as they could. After having done so, they were also told that their work was good but they did not get a reward. They received only this feedback before going on to the second activity.

Transfer task. After the children had completed the training task, a sheet filled with circles was presented to them. Gymnasts in both conditions received exactly the same instructions. They were asked to make as many pictures as possible and were told that one or more circles should be the main part of any picture. Five minutes were allowed for this task. Examples were already drawn on their sheets to ensure their understanding of the task. A set of extra circle sheets was available in case children needed them.

After the drawing session, children were asked to write down the topic below each picture they drew. For participants in the reward condition, the procedure ended at this moment and time was taken to debrief, thank, and answer questions for the children. In the no-reward condition, the notebooks and pencils were given during the debriefing, as a surprise to thank the girls for their participation.

Originality–Rarity Measure

The numerical measure of *originality* (or *statistical rarity*) used by Eisenberger and his colleagues (Eisenberger & Armeli, 1997; Eisenberger & Selbst, 1994) was used to assess both the themes and the picture topics. As in his studies, each product was given a numerical score. This was done by taking the reciprocal of each idea's frequency of occurrence (in the respective population) and multiplying it by 100. For example, if a theme was given by only 3 out of the 61 participants, the originality–rarity score for that theme was calculated as $3/61 \times 100 = 4.91$. Examples of original or rare themes were teeth, parents, and Backstreet Boys, and examples of common ones were comics, animals, and countries. Similarly, examples of rare pictures were a peach, a ring, and a doll, whereas common pictures were sun, balloon, and clock. Each participant was given two summary scores by calculating the mean of the originality scores for all their themes and all their pictures.

Judged Creativity Coding

In addition to Eisenberger's (1994) objective index of statistical rarity, Amabile's (1982) consensual method of judging creativity was used to assess the creativity of both the gala ideas and the drawings. This method requires judges to independently examine products and evaluate their creativity. Five undergraduates who were familiar with both the themes and pictures task examined a list of all the different themes suggested in the study and a similar list with all the pictures' topics. They were asked to read the entire list before giving ratings (for them to give comparative scores) and to use their own subjective definition of creativity and of appropriateness to rate the products. The judges were blind to the frequency of occurrence of the ideas listed. Next to each idea was a 5-point rating scale ranging from 1 (*low creativity*) to 5 (*high creativity*). Moreover, appropriateness of the themes was assessed with a 5-point rating scale ranging from 1 (*low appropriateness*) to 5 (*high appropriateness*). Because the drawing task was narrowly defined, it was thought that no appropriateness measure was necessary. There was no room to be inappropriate because a picture using a printed circle meets the demand. For each suggested product, the mean of the five judges' ratings was calculated. Summary scores of judged creativity were then calculated for each participant by taking their mean judged creativity score across all of their themes and pictures. Themes' score of appropriateness were calculated the same way. In the training task, examples of creative themes were the four elements, religions, and space. Themes judged as uncreative were gymnastics, school, and circus. In the transfer task, the most creative pictures were an astronaut, a mouse hole in a wall, and an electric plug. Finally, examples of uncreative pictures were a face, a ball, and a wheel.

Results

Preliminary Analysis

Reliability of ratings. The children produced a total of 155 different gala themes and 143 different picture ideas. Even though the judges were not specifically trained, they seemed to agree on the products' relative standing on both the creativity and the appropriateness scales. The highest reliability was obtained for creativity of pictures ($\alpha = .81$). The reliabilities for ratings of themes were lower but adequate (creativity judgments $\alpha = .71$; appropriateness judgments $\alpha = .67$).

Central Analysis

Analytic strategy. The four central dependent variables were originality (rarity) of themes and pictures, and the judged creativity of themes and pictures. We also examined the judged appropriateness of themes and the total number of themes and pictures produced by each participant. For all seven of these dependent variables we conducted 2 × 2 analyses of variance (ANOVAs) with group (no reward or reward) and age (younger or older) as between-subject factors. Age was converted to a categorical variable by a median-split procedure (*Mdn* age = 9). A preliminary set of analyses that included location of testing (Gym 1 vs. Gym 2) as a between-subject factor revealed no main effects or interactions for this variable. Hence, location was not included in the analyses presented here.

Training Task

Number of themes. The ANOVA revealed a significant main effect for age, $F(1, 53) = 21.71$, $p < .001$, as well as a marginally significant interaction between group and age, $F(1, 53) = 3.85$, $p = .06$. The main

effect indicated that older participants wrote more themes ($M = 7.57$) than the younger participants ($M = 4.03$). The interaction reflected the fact that in the younger group of children, rewards were associated with producing somewhat more themes than no rewards (reward $M = 5.14$; no reward $M = 2.92$), whereas among older children this pattern was reversed (reward $M = 7.93$; no reward $M = 7.20$).

Originality and judged creativity of themes. The ANOVAs revealed significant main effects for age on both measures: for originality, $F(1, 53) = 14.17, p < .001$, and for judged creativity, $F(1, 53) = 6.25, p = .02$. Surprisingly, the directions of these age effects were opposite for the two measures. As can be seen in Table 1, older participants produced themes that were less original (less rare) yet judged to be more creative than those produced by the younger participants. No other effects approached significance ($ps > .20$).

Appropriateness of themes. This analysis revealed a significant main effect for age, $F(1, 53) = 20.13, p < .001$. As can be seen in Table 2, older participants produced themes that were rated as much more appropriate than those of younger children. A marginally significant main effect for group was also revealed, $F(1, 53) = 3.45, p = .07$. It can be seen in Table 2 that participants wrote somewhat less appropriate themes in the reward condition than in the no-reward condition. Finally, a significant interaction between group and age level also emerged, $F(1, 53) = 4.33, p = .04$, showing that among younger participants, rewards were related to less appropriate themes, whereas no such pattern was evident among older participants.

Table 1. *Mean for Themes' Originality and Creativity by Age Groups*

Themes	Younger	Older
Rarity–Originality	42.4	22.8
Judged Creativity	2.65	2.94

Note: The first measure was calculated as the reciprocal of the frequency multiplied by 100, and the second was calculated on a scale ranging from 1 (*low creativity*) to 5 (*high creativity*).

Table 2. *Mean Ratings of Themes' Judged Appropriateness by Age and Reward*

Age	No Reward	Reward	Total
Younger	3.28	2.81	3.05
Older	3.56	3.56	3.56
Total	3.42	3.21	3.32

Note: Ratings made on a scale ranging from 1 (*low appropriateness*) to 5 (*high appropriateness*).

Transfer Task

Number of pictures. The 2 × 2 ANOVA revealed a significant main effect of age, $F(1, 53) = 11.78, p = .001$, showing that the older participants ($M = 8.43$) were more productive than younger ones ($M = 6.19$).

Originality and judged creativity of pictures. No effects approaching significance emerged for the measure of originality ($ps > .20$). For judged creativity, a significant main effect for age was obtained, $F(1, 53) = 4.57, p = .04$, indicating that older children drew more creative pictures ($M = 2.27$) than the younger children ($M = 2.05$). The ANOVA also yielded a marginally significant main effect of group, $F(1, 53) = 2.74, p = .10$. Drawings made in the no-reward group were judged to be somewhat more creative than the drawings made by the rewarded children ($Ms = 2.25$ and 2.08, respectively).

Supplemental Analysis

A potential problem with our research design was that some children were rewarded for producing relatively few themes during the training task. Eisenberger and Cameron (1996) argued that rewards would only enhance creativity when they were administered contingent on the display of high divergent thinking. To examine whether the number of themes produced during the training period moderated the impact of rewards on originality and judged creativity, we conducted a 2 × 2 ANOVA with group and number of themes (low or high) as between-subject factors. These analyses revealed nearly identical results to those reported earlier. There were no effects approaching significance for originality and there was only a marginally significant main effect for group on the

measure of judged creativity ($p = .07$). More important, there was no evidence that the number of themes produced (i.e., the level of divergent thinking displayed on the training task) influenced the impact of rewards on the originality or creativity of the transfer task.[1]

Discussion

In this study, we examined whether expected rewards influenced creativity in (a) a training task (generating themes for a gymnastics gala), in which a reward contingency was either present or absent, and (b) the following transfer task (using circles to make drawings). Although our tasks were similar to those used by Eisenberger and Selbst (1994), the effects we found were in the opposite direction of those previously obtained.

In their study, Eisenberger and Selbst (1994) did not find group differences in the training phase. Similarly, we found no group difference in creativity by using either his originality (rarity) measure or the consensual measure of judged creativity. However, because our sample had a wide age range, we could explore relations between age and creativity. We found opposite patterns for the two different creativity measures. Although younger participants wrote more rare or original themes than older participants, they produced themes judged as less creative than themes given by older participants. Which measure is more valid? Does creativity increase or decrease with age?

Torrance (1962) described the general pattern of the developmental curve for creativity as a steady increase from ages 5 to 18, interrupted by two drops of short duration, one around age 10, and another around age 12. The age pattern obtained for the judged creativity measure was consistent with this typical growth curve. By contrast, the originality measure for themes, based on rarity of occurrence, contradicted the common increase of creativity with age. This surprising set of opposing findings shows how the two methods do not assess the same thing. Perhaps the first one measures divergence instead of creativity. This would be consistent with the idea that although divergent thinking can be a good estimate of the potential for creative thinking, it is not synonymous with creativity (Runco & Chand, 1995) and "it has been recognized for some time that creative ideas are not just original, but are in fact original *and* appropriate" (Runco & Chand, 1995, pp. 255–256).

In addition to the creative dimension, the appropriateness of the themes was also assessed. The task of finding gala themes is a much more open one than the pictures made with circles and it resulted in all kinds of ideas. Some were rare but not at all suitable for a gymnastics show (e.g., teeth, given names, and the hot or cold water). The younger participants' suggestions were found to be less appropriate than the suggestions of the older participants. Thus, as they were writing rarer themes, they also wrote less appropriate ones. This age effect on appropriateness confirms the originality measure's potential problem with capturing oddness rather than creativity. Interestingly, Hennessey (1998) made a similar comment about the measure described in Eisenberger and Cameron's meta-analysis (1996):

> Creativity was operationalized as the simple statistical infrequency of responses. Although this simple measure can legitimately be termed *originality* or *divergent thinking*, it does not adequately capture the elements of creativity as it is generally defined in the literature: novelty combined with appropriateness, value, usefulness. (Hennessey, 1998, p. 674)

Therefore, we suggest that although judged creativity assesses what people generally think of as a creative product, the rarity–originality calculation isolates the novelty aspect to the detriment of appropriateness. Although the latter measure is more objective, we think that it does not really get at what creativity is. Amabile's (1983) measure relies on a subjective criteria for creativity but it captures what "people can recognize and often agree on, even if they are not given a guiding definition" (p. 360).

It is easy to understand that younger children have less appropriate ideas. However, the significant inter-

[1] A 2 × 2 × 2 × 2 repeated measures multivariate analysis of variance (MANOVA) was also performed in which type of measure (judged creativity or rarity) and type of task (themes or pictures) served as within-subject factors and age (young or old) and group (reward or no reward) served as between-subject factors. Creativity and rarity scores were standardized prior to this analysis. The results revealed a significant main effect for age, $F(1, 53) = 13.23, p < .001$, indicating that older participants generally scored higher on both rarity and creativity across both tasks. This main effect was qualified by an Age × Measure interaction, $F(1, 53) = 15.02, p < .001$, and an Age × Task interaction. These interactions reflected the fact that it was especially on the judged creativity measures and the themes task that older children scored higher than younger children. There were no effects for group in the MANOVA. However, we believe the specific ANOVAs were more precise tests of our hypotheses regarding the impact of rewards.

action shows that themes given by young participants were less appropriate in the rewarded group than in the other group. The reward contingency was probably adding pressure on the young children and led them to meet the demand at all cost. This seems to be the case because the younger children wrote significantly more themes in the reward condition. The pressure to write as many themes as possible was effective regarding the productivity but the cost seems to have been a lesser quality of product.

The major result of interest concerns the transfer task. Whereas behaviorists argue that rewards for divergent thinking in a task increase creativity in the subsequent task, cognitively oriented psychologists predict that the expectancy of rewards inhibits creativity in later tasks. Our study provided modest support for the latter point of view. As expected, reward was associated with somewhat lower judged creativity in the transfer task. The drawings made in the reward group were judged as somewhat less creative than the drawings made by children who did not expect any external reward. Many other studies have also confirmed that expected rewards have a detrimental effect on creativity (Amabile, 1996; Condry, 1977).

Our study had some limitations. First, the sample was small and perhaps unrepresentative of all children because the participants were female gymnasts. However, the tasks did not require any special skills and they could easily be replicated with any population of children. Second, although appropriateness is seen to be critical in creativity, its assessment is difficult. The interrater reliability coefficient (α = .67) was lower for this dimension than for judged creativity. This problem of subjectivity was already underscored by Runco and Charles (1993). Finally, our training task did not demand a specific number of themes. Although the demand for high divergent thinking was present in the instructions ("as many as possible"), some children may have been rewarded for a small number of ideas. This would reflect a reward for low divergent thinking, shown to be detrimental by Eisenberger and Selbst (1994). However, supplemental analyses that controlled for number of themes generated (and hence divergent thinking) indicated that the negative impact of rewards on judged creativity remained marginally significant.

These results would seem to have implications for how to encourage creativity in children. External rewards are known to increase the quantity of behavior, but whether they facilitate the quality of productive behavior is still debatable. Creativity represents a quality-based dimension that may not follow the reinforcement rule. Parents, teachers, and coaches should be aware that promising an external payoff for a task may decrease children's creativity at a subsequent task. The detrimental effect probably comes from extrinsic motivation that leads children to feel controlled. Although it is natural for socializing agents to think that promising rewards will enhance children's motivation and performance (Boggiano, Barrett, & Weiher, 1987), there is now considerable evidence that not only is this assumption false but also that such incentives may backfire and undermine both intrinsic motivation and creativity (Amabile, 1996; Ryan, Mims, & Koestner, 1983). Future research needs to consider exactly which aspects of the creative processes (e.g., problem finding, ideation, evaluation) are influenced by motivational factors (Runco & Chand, 1995).

References

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43,* 997–1013.

Amabile, T. M. (1983). Social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology, 45,* 357–376.

Amabile, T. M. (1996). *Creativity in context.* Boulder, CO: Westview.

Amabile, T. M., Goldfarb, P., & Brackfield, S. C. (1990). Social influences on creativity: Evaluation, coaction and surveillance. *Creativity Research Journal, 3,* 6–21.

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.

Barron, F. (1968). *Creativity and personal freedom.* New York: Van Nostrand.

Boggiano, A. K., Barrett, M., & Weiher, A. W. (1987). Use of maximal operant principle to motivate children's intrinsic interest. *Journal of Personality and Social Psychology, 53,* 866–879.

Condry, J. (1977) . Enemies of exploration: Self-initiated versus other-initiated learning. *Journal of Personality and Social Psychology, 35,* 459–477.

Deci, E. L., & Ryan, R. M. (1987). Supportive autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53,* 1034–1047.

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

Eisenberger, R., & Armeli, S. (1997) . Can salient reward increase creative performance without reducing intrinsic creative inter-

est? *Journal of Personality and Social Psychology, 72,* 652–663.

Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist, 51,* 1153–1166.

Eisenberger, R., & Selbst, M. (1994). Does reward increase or decrease creativity? *Journal of Personality and Social Psychology, 66,* 1116–1127.

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

Koestner, R., & McClelland, D. C. (1990). Perspectives on competence motivation. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 527–548). New York: Guilford.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology, 28,* 129–137.

McGraw, K. O., & McCullers, J. C. (1979). Evidence of a detrimental effect of extrinsic incentives on breaking mental set. *Journal of Experimental Social Psychology, 15,* 285–294.

Okuda, S. M., Runco, M. A., & Berger, D. E. (1991). Creativity and the finding and solving of real-world problems. *Journal of Psychoeducational Assessment, 9,* 45–53.

Runco, M. A. (1996). Personal creativity: Definition and developmental issues. In M. A. Runco (Ed.), *New directions for child development* (Vol. 72, pp. 3–25). San Francisco: Jossey-Bass.

Runco, M. A., & Chand, I. (1995). Cognition and creativity. *Educational Psychology Review, 7,* 243–267.

Runco, M. A., & Charles, R. (1993). Judgements of originality and appropriateness as predictors of creativity. *Personality and Individual Differences, 15,* 537–546.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology, 45,* 736–750.

Schwartz, B. (1982). Reinforcement-induced behavioral stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General, 111,* 23–59.

Torrance, E. P. (1962). *Guiding creative talent.* Englewood Cliffs, NJ: Prentice Hall.

Urban, K. K. (1991). On the development of creativity in children. *Creativity Research Journal, 4,* 177–191.