
Gender Differences in the Effects of Extrinsic Motivation on Creativity*

ABSTRACT

Four studies were conducted to assess gender differences in the effects of extrinsic motivation on creativity. The first study replicated an earlier one (Baer, 1997) in which expectation of evaluation lowered the creativity of middle school girls, but not boys. The second study investigated the effects of doing work for reward; again, middle school girls' creativity suffered, but not boys'. The third study, also with middle school subjects, investigated the impact of expecting ungraded feedback; this reduced both the overall negative impact of expecting evaluation and the gender difference in this regard. The fourth study investigated the impact of evaluation expectation on second-grade subjects and found that boys', but not girls', creativity was increased when they expected evaluation.

GENDER
DIFFERENCES IN
THE EFFECTS OF
EXTRINSIC
MOTIVATION
ON CREATIVITY

Thirteen years after her seminal publication, *The Social Psychology of Creativity* (Amabile, 1983), Amabile (1996) and her colleagues continue to produce studies that demonstrate clearly the negative effects of extrinsic motivation on creativity. Conditions that typically lead to a decrease in creative performance include expecting that one's work will be evaluated and undertaking a task in order to receive a reward, both of which are common conditions in both schools and workplaces.

Amabile's (1996) work is rooted in the "overjustification" hypothesis (Bem, 1972; deCharms, 1968; Kelly, 1967, 1973; Lepper, Greene & Nisbett, 1973), which states that, under certain conditions, extrinsic constraints may lead to a decrease in intrinsic motivation. This hypothesis has successfully predicted

* This research was supported, in part, by a research grant from Rider University.

a variety of research results (e.g., Calder & Staw, 1975; Deci, 1971; Lepper & Greene, 1975; Ross, 1975), many of them decidedly counterintuitive (e.g., Lepper, Greene & Nisbett, 1973). These investigations have sometimes been called "hidden cost of reward" research (Lepper & Greene, 1978) because they demonstrate an often undocumented negative effect — in the form of a loss of intrinsic motivation — as an unintended result of reinforcing a desired behavior.

CHANGES IN THE
INTRINSIC
MOTIVATION
PRINCIPLE OF
CREATIVITY

Although evidence continues to accumulate to support the importance of intrinsic motivation in creativity and the detrimental impact of extrinsic constraints, there have been some changes in the theory. Amabile's (1996) current "Intrinsic Motivation Principle of Creativity" is more careful and circumscribed than her earlier "Intrinsic Motivation Hypothesis" (Amabile, 1983). The earlier version stated without qualification that "the intrinsically motivated state is conducive to creativity, whereas the extrinsic state is detrimental" (p. 91). The new version adds that "informational or enabling extrinsic motivation can be conducive, particularly if initial levels of intrinsic motivation are high" (p. 119).

It is helpful to consider briefly some of the research and theory that have led Amabile (1996) to qualify her earlier hypothesis. Below are listed six conditions under which the effects of evaluation and rewards on creativity may be more complex than the original hypothesis suggested.

1. Some extrinsic constraints can have a positive impact on creative performance. Building on Deci and Ryan's (1985) distinction between "controlling" and "informational" extrinsic motivators, Amabile (1996) argued that "informative, constructive feedback and evaluation" (p. 152) can have a positive influence on creativity. This claim is supported by three real-world (nonexperimental) studies of work environments using interview and questionnaire measures.
2. Receiving an unexpected reward is apparently quite different from undertaking a task in order to earn a reward, and such unexpected rewards can have exactly the opposite effect on creativity of contracted-for rewards (Amabile, Hennessey, & Grossman, 1986).
3. Personality variables such as shyness may mediate the effects of extrinsic constraints. For example, in a study by Cheek and Stahl (1986), there was no main effect of

evaluation expectation on poem-writing creativity; however, a separate analysis revealed that shy subjects were significantly less creative than subjects who were not shy when they expected that their poems would be evaluated.

4. Subjects who have been trained to distance themselves cognitively from extrinsic constraints and to focus on intrinsic motivation can limit the negative impact of extrinsic constraints, and in some cases children "immunized" from the effects of extrinsic constraints through training actually perform more creatively when offered rewards than when not offered rewards (Hennessey, Amabile, & Martinage, 1989; Hennessey & Zbikowski, 1993).
5. Subjects with different levels of skill in a task domain may respond differently to extrinsic constraints. In a study by Conti and Amabile (1995) which assessed the creativity of computer programs created by subjects whose programming skills had been assessed using a separate measure, high-skill subjects performed less creatively when they anticipated evaluation, but the reverse was true of low-skill subjects, whose programs were more creative when they expected evaluation.
6. Rewards that confirm competence or which make it possible for an individual to do interesting work may serve as what Amabile (1996) termed "synergistic extrinsic motivators" (p. 118). These rewards do not undermine one's sense of self-determination and can lead to higher, rather than lower, creative performance.

Despite these qualifications, research (e.g., Amabile, 1985; Amabile, Goldfarb, & Brackfield, 1990; Amabile, Hennessey, & Grossman, 1986; Hennessey & Amabile, 1988) continues to support the general conclusion that as one's motivation tends to become more extrinsic and less intrinsic, one's creative performance is decreased. Extrinsic motivation in the form of anticipating that one will receive either an evaluation or a reward for one's work has generally resulted in a decrease in creative performance in a wide variety of studies, although some recent studies reported by Amabile (1996) have been inconclusive regarding the effects of task motivation on creative performance. This generally negative impact of extrinsic motivational constraints may be especially significant in

educational settings, where rewards are frequently employed, where evaluating students' work is often necessary, and where neither students' levels of skill at tasks they are undertaking nor their initial levels of intrinsic motivation in those tasks are likely to be high.

GENDER
DIFFERENCES IN
CREATIVITY

Many of Amabile's (1983) earlier studies employed only women. In her initial work with collage-making she found that women tended to produce more creative collages than men, and to avoid possible complicating effects, she decided to use only female subjects in most studies. Amabile (1996) has since abandoned this practice, in part because she "did not wish to develop a social psychology of female creativity" (p. 78) and in part because gender differences in creative performance have tended, in fact, to be rare.

Gender differences in creativity are indeed uncommon in both divergent-thinking testing and in consensual assessment of creative products (Baer, in press). Two recent studies, however, suggest that motivational constraints may, under some conditions, impact girls and boys differently. That is, there may in fact be differences between a "social psychology of female creativity" and a "social psychology of male creativity." In one recent study, Baer (1997) had eighth-grade girls and boys write original stories and poems under conditions in which evaluation was either expected (and highly salient) or unexpected. Creativity ratings of the poems and stories yielded a significant gender x motivational condition effect; girls' creativity decreased markedly when evaluation was anticipated, but boys' creativity did not. And in an unpublished study by Conti, Collins, and Picariello (1995; cited by Amabile, 1996) in which boys and girls made collages in competitive and noncompetitive situations, competition lowered girls' creativity, but not boys'. In fact, "it appears that boys may show an even higher level of creativity on this artistic activity under competitive conditions" (Amabile, 1996, p. 240).

OVERVIEW OF
STUDIES 1-4

The four studies reported below were designed to:

1. Replicate Baer's (1997) study of the differential effects of anticipated evaluation on the creativity of middle school girls and boys using a collage-making activity rather than poetry- and story-writing;
2. Investigate the differential effects of contracted-for reward on the creativity of middle school girls and boys;

3. Investigate the differential effects of nongraded, formative feedback (as opposed to the summative, high-stakes evaluation of Study 1) on the creativity of middle school girls and boys; and
4. Investigate the differential effects of anticipated evaluation on the creativity of second-grade girls and boys.

STUDY 1:
DIFFERENTIAL
EFFECTS OF
ANTICIPATED
EVALUATION ON
THE CREATIVITY OF
MIDDLE SCHOOL
GIRLS' AND BOYS'
COLLAGES

The goals of this study were (a) to replicate a previous study (Baer, 1997) which found a significant gender difference in the effects of anticipated evaluation on the creativity of middle school children and (b) to do so using a task from a different domain than that investigated in the previous study. The latter goal is based on research (Baer, 1991, 1993, 1994) suggesting that creativity may be task-specific and on the more general objective of seeing how robust the results of the previous study might be.

METHOD
Subjects

The subjects were 70 seventh- and eighth-grade students, 35 girls and 35 boys, attending public school in a suburban New Jersey district. All students made collages in their regular art classes. There were a total of four art classes, all taught by the same teacher. Although there was some ability grouping in this school for instructional purposes, art classes were heterogeneously mixed. The assignment of two classes to each condition (experimental and control) was done randomly. All students in each of the classes participated. It should be noted that there was no overlap of subjects in Studies 1-4 and no subjects were informed in any of the studies that other groups were making collages under different conditions.

Tasks

Each subject was given a 14" X 22" piece of white tagboard, a bottle of glue, and a set of over one hundred pre-cut construction paper designs (for example, hearts, butterflies, squares, circles, and triangles) and asked to make an "interesting, imaginative design." The materials each student received were identical. Subjects had 47 minutes (the length of the class period) to complete their collages.

The collages were later rated for creativity by four art educators. Raters were given the following instructions: "There is only one criterion in rating these collages: creativity. I realize that creativity doesn't exist in a vacuum, and to some extent creativity probably overlaps other criteria one might apply — aesthetic appeal, organization, use of color, novelty, complexity, balance, symmetry, technical goodness, neatness, or detail, for example — but I ask you to rate the collages solely on

the basis of your thoughtful-but-subjective opinions of their creativity. The point is, you are the expert, and you needn't defend your choices or articulate a definition of creativity. What creativity means to you can remain a mystery—what I want you to do is use that mysterious expert sense to rate the collages for creativity.”

A 1.0 (low creativity) to 5.0 (high creativity) rating scale was used, and judges were encouraged to use the full scale; that is, they were encouraged not to concentrate their ratings around a single score point. There were no specific limitations about the numbers of collages that were to fall in each scoring range, however. The mean rating of all judges was used as the creativity rating of each collage.

The raters worked independently of one another; they did not know the students, their gender, or the specific purpose of the study; they rated the collages in different random orders; and they were paid for their work. The coefficient alpha inter-rater reliability was .80.

This consensual technique of assessing the creativity of collages has been validated extensively by Amabile (1982, 1996). Although she initially used the Spearman-Brown prediction formula to calculate inter-rater reliability, Amabile (1996) has in recent years used coefficient alpha. Following this model, coefficient alpha was employed in all four investigations reported below (Studies 1-4).

Procedure

Two of the four classes were randomly assigned to be the Expecting Evaluation group and two to be the Not Expecting Evaluation group. The experimenter went to all four classes and directed the collage-making activity. In both the Expecting Evaluation and Not Expecting Evaluation groups the basic collage-making instructions were the same. The difference was in how the purpose of the activity was explained.

In the Not Expecting Evaluation classes, the experimenter explained that he was there to try out a new task that he might use in some experiments in the future. Students were encouraged to make the most interesting collages they could. They were not asked to put their names on their collages; they were not observed closely as they worked; and there was no indication that there would be any kind of evaluation made of their collages. They were encouraged to enjoy the activity, which was described as a “fun collage-making activity.”

In the Expecting Evaluation classes, the experimenter explained that he was from the New Jersey State Department of Education and was responsible for assessing the artistic

ability of students in the state. The collage-making task, he explained, was being used to evaluate the artistic ability of middle school students. Several art educators would assess the quality of the collages the students made and the result of their assessment would be sent to the students' school. Students were encouraged to make the most interesting collages they could. They were asked to put their names on the back of the tagboard on which they would construct their collages before they began work.

In this and the three studies reported below, students were told at a later date the truth about the experimental procedure and the basic intrinsic motivation theory of creativity that guided their conception.

Results A 2x2 ANOVA was performed. The primary hypothesis was that there would be a gender x motivational condition interaction effect. This prediction was confirmed. Neither gender nor condition was statistically significant. Full results of the analysis of variance are reported in Table 1; group means are reported in Table 2.

TABLE 1. Study 1: Evaluation/No Evaluation, 7th & 8th Grade (Analysis of Variance).

Source	<i>F</i>	<i>df</i>	<i>p</i>
Gender	0.001	1, 66	0.976
Condition	3.236	1, 66	0.077
Gender x Condition	4.693	1, 66	0.034

TABLE 2. Study 1: Evaluation/No Evaluation, 7th & 8th Grade (Comparisons of Means).

Comparison	Group	Mean	<i>SD</i>
Gender x Condition	Female, Evaluation	2.487	0.927
	Female, No Evaluation	3.316	0.825
	Male, Evaluation	2.934	0.710
	Male, No Evaluation	2.857	0.931

A separate analysis of only the creativity ratings of the collages made by the 35 girls yielded a significant effect for motivational condition ($F(1,33) = 7.773, p = .009$).

Discussion of
Study 1

This study replicates Baer's (1997) finding that the creativity of middle school girls was diminished when they expected their work would be evaluated and extends that research finding to a different task domain. (The previous work used poetry- and story-writing rather than collage-making.)

It is interesting that the overall effect of task motivation was not statistically significant, but that a separate analysis of just the girls in the study yielded a significant effect for task motivation. The latter result, but not the former, is in line with Amabile's (1996) general findings. This partial contradiction is somewhat resolved when one recalls that many of the studies of the impact of evaluation expectation on creative performance upon which Amabile based her conclusions were conducted only with female subjects.

STUDY 2:
DIFFERENTIAL EFFECTS
OF REWARD ON THE
CREATIVITY OF MIDDLE
SCHOOL GIRLS' AND
BOYS' COLLAGES

The primary goal of this study was to learn if the gender differences observed in Study 1, which found differential effects of one kind of extrinsic constraint (evaluation expectation), would extend to the other most common kind of extrinsic constraint, reward.

METHOD
Subjects

The subjects were 49 eighth-grade students, 23 girls and 26 boys, attending public school in a suburban New Jersey district. All students made collages in their regular art classes. There were a total of four art classes, all taught by the same teacher. Although there was some ability grouping in this school for instructional purposes, art classes were heterogeneously mixed. The assignment of two classes to each condition (experimental and control) was done randomly. With one exception (explained in the Procedure section), all students in each of the classes participated.

Tasks

This was identical to the task used in Study 1. There were four raters. The coefficient alpha inter-rater reliability was .85.

Procedure

Two of the four classes were randomly assigned to be the Reward group and two to be the No Reward group. The experimenter went to all four classes and directed the collage-making activity. In both the Reward and No Reward groups the basic collage-making instructions were the same, and with one difference were the same as the instructions given to the Not Expecting Evaluation group in Study 1. That is, all subjects were told that they were being asked to help the experimenter try out a new task that might be used in future research. All students were encouraged to make the most interesting collages they could. They were not asked to put their names

on their collages; they were not observed closely as they worked; and there was no indication that there would be any kind of evaluation made of their collages. They were encouraged to enjoy the activity, which was described as a "fun collage-making activity."

The No Reward group was given no choice regarding participation and received no reward other than the experimenter's expression of appreciation at the conclusion of the activity. The Reward groups were given a choice whether or not to participate in the activity and were offered unspecified bonus points toward their art grade for participation. Students were told that if they chose not to participate, they could spend the class period reading or doing other work quietly at their desks. Only one student chose not to participate. It is unlikely that this limited degree of self-selection influenced the outcome of the study; it does suggest, however, that the students did feel the choice was a real one, and that they were participating in order to earn the reward.

Results A 2x2 ANOVA was performed. The primary hypothesis was that there would be a gender x motivational condition interaction effect. This prediction was confirmed. Motivational condition was statistically significant in the direction predicted by previous research (i.e., the No Reward group was more creative). Full results of the analysis of variance are reported in Table 3; group means are reported in Table 4.

TABLE 3. Study 2: Reward/No Reward, 8th Grade (Analysis of Variance).

Source	<i>F</i>	<i>df</i>	<i>p</i>
Gender	0.043	1, 45	0.837
Condition	10.943	1, 45	0.002
Gender x Condition	5.491	1, 45	0.024

TABLE 4. Study 2: Reward/No Reward, 8th Grade (Comparisons of Means).

Comparison	Group	Mean	<i>SD</i>
Gender x Condition	Female, Reward	2.142	0.813
	Female, No Reward	3.527	0.833
	Male, Reward	2.665	0.873
	Male, No Reward	2.902	0.895

A separate analysis of only the creativity ratings of the collages made by the 23 girls yielded a significant effect for motivational condition ($F(1,21) = 16.275, p = .001$). Looking only at the creativity ratings of the collages made by the 26 boys, there was no effect for motivational condition ($F(1,24) = 0.465, p = .502$).

Discussion of
Study 2

This study is consistent with previous research (Amabile, 1996) regarding the effects of contract-for reward. It adds a new dimension to that research by suggesting that the primary impact of rewards on creative performance among middle school students is among girls, with a smaller overall (mixed girls and boys) effect and no effect among boys. This is consistent with the findings of Study 1 of the differential impact of anticipated evaluation on boys and girls.

STUDY 3:
DIFFERENTIAL EFFECTS
OF ANTICIPATED
UNGRADED
FEEDBACK ON THE
CREATIVITY OF
MIDDLE SCHOOL
GIRLS' AND BOYS'
COLLAGES

The primary idea that motivated this study was to see if anticipation of "informative, constructive feedback" (Amabile, 1996, p. 152) rather than high-stakes, summative evaluation (as in Study 1) might lessen the negative impact of anticipated evaluation on the creativity of middle schools girls. A secondary goal was to compare the impact of anticipating such feedback on middle school girls and boys.

METHOD
Subjects

The subjects were 60 seventh-grade students, 27 girls and 33 boys, attending public school in a suburban New Jersey district. All students made collages in their regular art classes. There were a total of four art classes, all taught by the same teacher. Although there was some ability grouping in this school for instructional purposes, art classes were heterogeneously mixed. The assignment of two classes to each condition (experimental and control) was done randomly. All students in each of the classes participated.

Tasks

This was identical to the task used in Study 1. There were four raters. The coefficient alpha inter-rater reliability was .87.

Procedure

Two of the four classes were randomly assigned to be the Expecting Feedback group and two to be the Not Expecting Feedback group. The experimenter went to all four classes and directed the collage-making activity. In both the Expecting Feedback and Not Expecting Feedback groups the basic collage-making instructions were the same. The difference was in how the purpose of the activity was explained.

The Not Expecting Feedback classes received the same treatment as the Not Expecting Evaluation group in Study 1. That is, the experimenter explained that he was there to try

out a new task that he might use in some experiments in the future. Students were encouraged to make the most interesting collages they could. They were not asked to put their names on their collages; they were not observed closely as they worked; and there was no indication that there would be any kind of evaluation made of their collages. They were encouraged to enjoy the activity, which was described as a "fun collage-making activity."

In the Expecting Feedback classes, the experimenter explained that he was an art educator who had developed a tool for helping give students feedback on their strengths as artists. He explained that no overall evaluation or grade would be made of their collages, but that he would be able to provide them with feedback about specific art abilities and guidance for how to continue to improve their artistic skills based on how they did their collages. He responded to questions for more details about the nature of the feedback by assuring them that the feedback would be positive in tone because he would be looking for strengths rather than weaknesses; however, he couldn't say more before they actually made their collages for fear of inadvertently influencing the way they made their collages. Students were encouraged to make the most interesting collages they could. They were asked to put their names on the back of the tagboard on which they would construct their collages before they began work.

Results

Looking first at the effect of motivational condition on the 27 girls in the study, the results ($F(1,25) = 4.215, p = .051$) border on statistical significance. The mean difference between the two groups' scores is .706 points on a four-point (1.0-5.0) scale. Although one might argue in this case for flexibility in applying the conventional .05 standard (the observed .051 two-tailed p -value would, for example, pass Abelson's (1995) "lopsided test" [p. 59] and follow his recommendation that the .05 convention be employed "with less than total rigidity" [p. 131]), readers may differ in the liberality of their interpretations. As in previous studies, anticipation of evaluation, even in the form of constructive, positive feedback, appeared to decrease the creative performance of the girls. For the 33 boys there was no effect ($F(1,31) = 0.017, p = .896$), nor was there any overall effect of motivational condition.

A 2x2 gender x motivational condition ANOVA revealed no significant effects. Full results of this analysis of variance are reported in Table 5; group means are reported in Table 6.

TABLE 5. Study 3: Feedback/No Feedback, 7th Grade (Analysis of Variance).

Source	<i>F</i>	<i>df</i>	<i>p</i>
Gender	0.457	1, 56	0.502
Condition	1.440	1, 56	0.235
Gender x Condition	1.955	1, 56	0.168

TABLE 6. Study 3: Feedback/No Feedback, 7th Grade (Comparisons of Means).

Comparison	Group	Mean	<i>SD</i>
Gender x Condition	Female, Feedback	2.292	0.703
	Female, No Feedback	2.998	1.038
	Male, Feedback	2.856	1.064
	Male, No Feedback	2.802	1.230

DISCUSSION OF
STUDY 3

Interpretation of these results is somewhat tricky. It appears that girls' creativity was influenced negatively by the anticipation of feedback, but this effect did not quite reach statistical significance. It should be noted that this experimental study is somewhat different from the questionnaire and interview studies upon which Amabile (1996) based her conclusion that informative, constructive feedback might not negatively impact creativity (and might, she argued, even support creative performance). An unknown experimenter coming into a middle school classroom offering to provide constructive feedback is not the same as a trusted teacher or mentor with a track record of providing supportive, nonthreatening feedback. Also, what was manipulated in this study was expected constructive feedback, not actual feedback. Overall, these results are not inconsistent with Amabile's suggestion that on-going, constructive feedback need not diminish creative performance. Once again, however, it appears that whatever negative impact on creative performance there might be from anticipated evaluation — even in the form of constructive, ungraded feedback — is limited to girls.

STUDY 4:
DIFFERENTIAL EFFECTS
OF ANTICIPATED
EVALUATION ON THE
CREATIVITY OF
SECOND-GRADE
GIRLS' AND BOYS'
COLLAGES

The goal of this study was to begin to investigate the development of the gender differences observed in previous studies (Studies 1-3 above; Baer, 1997; Conti, Collins, & Picariello, 1995) of the impact of extrinsic motivation on creative performance. This study was similar to Study 1 except that second-grade students were the subjects rather than middle school students.

METHOD

Subjects

The subjects were 81 second-grade students, 41 girls and 40 boys, attending four different public schools in a suburban New Jersey district. All students made collages in their regular classes. Unlike Studies 1-3, in which all students were in classes with the same art teacher, each class of students had a different teacher. There was no control in this study for possible differences in classroom climates and expectations other than random assignment of two classes to each condition (experimental and control). All students in each of the classes participated.

Tasks

This was identical to the task used in Study 1. There were nine raters. The coefficient alpha inter-rater reliability was .91.

Procedure

Student teachers led the collage-making activities in all classes, which were assigned randomly to condition. In two of the classes, evaluation was emphasized; students in these Expecting Evaluation classes were told that their collages would be evaluated, and that these evaluations were important to the student teacher because her supervisor would consider the students' success in making collages as part of his evaluation of the student teacher's success as a teacher. This condition is similar to one used in an unpublished study by Berglas, Amabile, and Handel (1981; reported in Amabile, 1996) in which student teacher evaluations were used to make evaluation more salient with young children, with the exception that rather than using prior evaluation as the cue to lead students to anticipate evaluation of their work, students were simply told that their collages would be evaluated. In the other two classes the activity was simply introduced as a fun collage-making activity with no mention of evaluation. As art work was not typically evaluated in either of these Not Expecting Evaluation classrooms, it is unlikely that the students anticipated their collages would be evaluated.

Results

A 2x2 ANOVA revealed no significant effects. Full results of this analysis of variance are reported in Table 7; group means are reported in Table 8. Although the gender x motivational condition effect was not statistically significant, separate analy-

TABLE 7. Study 4: Evaluation/No Evaluation, 2nd Grade (Analysis of Variance).

Source	<i>F</i>	<i>df</i>	<i>p</i>
Gender	1.511	1, 77	0.223
Condition	3.052	1, 77	0.085
Gender x Condition	1.607	1, 77	0.209

TABLE 8. Study 4: Evaluation/No Evaluation, 2nd Grade (Comparisons of Means).

Comparison	Group	Mean	<i>SD</i>
Gender x Condition	Female, Evaluation	2.645	0.715
	Female, No Evaluation	2.733	0.861
	Male, Evaluation	3.190	0.954
	Male, No Evaluation	2.638	0.710

ses of the creativity ratings given to the boys' and girls' collages revealed a *positive* impact of anticipated evaluation on the boys that was marginally significant statistically ($F(1,38) = 4.142, p = .049$).

Discussion of
Study 4

These results are somewhat inconclusive. It is interesting that the second-grade boys in this study appeared to be *more* creative when they anticipated evaluation. Perhaps this is related to the finding of Conti and Amabile (1995) that low-skill subjects were more creative when they expected evaluation. One possible interpretation of both the Conti and Amabile finding and the results of Study 4 is that the expectation of evaluation led to greater effort among subjects who, with little skill and perhaps little intrinsic motivation, might have made little effort without the extrinsic task constraint of anticipated evaluation.

These results do not provide evidence that the negative impact of anticipated evaluation on girls' creativity demonstrated in Studies 1-3 and elsewhere (Baer, 1997; Conti, Collins, & Picariello, 1995) is rooted in the behavior of much younger girls. Further research is needed to trace the development of the influence of extrinsic constraints (such as the expectation of evaluation) on girls' creativity, but Study 4 suggests that this development may occur sometime *after* second grade.

GENERAL
DISCUSSION

The effects of evaluation and rewards on creativity and intrinsic motivation have been well documented (e.g., Amabile, 1985, 1996; Amabile, Goldfarb, & Brackfield, 1990; Amabile, Hennessey, & Grossman, 1986; Berglas, Amabile, & Handel, 1981; Hennessey & Amabile, 1988; Lepper & Greene, 1978). It has been assumed that these effects were shared equally among boys and girls and women and men; in fact, Amabile (1996), after using primarily female subjects for many years, has more recently reversed this practice in favor of using subjects of both genders.

The results of Studies 1-3, together with previous work by Baer (1997) and Conti, Collins, and Picariello (1995), suggest that, at least for middle school-aged children, it may be necessary to develop somewhat different social psychologies for girls and boys in regard to the impact of extrinsic constraints on creativity. Expectation of evaluation and undertaking a task for reward appears to have a significantly different impact on girls and boys.

How might the observed differences between the impact of extrinsic motivational constraints on the creativity of early adolescent girls and boys best be interpreted? There are at least three viable interpretations of the observed gender x motivational condition difference:

1. Early adolescent girls may be *more attentive* than boys to cues from their social environment regarding motivational constraints in the forms of evaluations and rewards and how such constraints should determine the appropriate motivational set for a task. In consequence, the effects of task constraints favoring extrinsic motivation would have greater impact on their creative performance.
2. Early adolescent girls may simply *respond more* to differences in their own motivational set than do boys. As a result, even when both girls and boys experience the same levels of intrinsic and extrinsic motivation, these different kinds of motivation would have greater impact on girls than on boys. Under this interpretation, girls and boys would be equally receptive to cues from their social environment regarding the appropriate type of motivation; however, the difference between intrinsic and extrinsic sources of motivation would be more significant for girls than boys in terms of their effects on creative performance.

3. Early adolescent girls and boys may *respond differently* to cues from their social environment regarding the appropriate type of motivation for a given task. For example, although both girls and boys may be equally attentive to cues about anticipated evaluation or rewards, such anticipation may lead to a decrease in intrinsic motivation for girls but not for boys. Extrinsic constraints might even, under some conditions, have opposite effects on boys and girls.

The results of Studies 1-3 do not help us choose among these interpretations, and each could equally well account for all observed differences. The first interpretation is consistent with research suggesting that girls at this age are typically more attentive to interpersonal communications in general and the expectations of others in particular (Gilligan, Lyons, & Hanmer, 1990; Pool, 1994). Even if this difference in attentiveness does result in different levels of intrinsic or extrinsic motivation for girls and boys, however, it is still possible that girls and boys might respond differently to the cues to which they *do* attend; that is, the three interpretations are not mutually exclusive (in fact, one could argue that Interpretation 2 is actually just a subcategory of Interpretation 3), and two or even all three factors could be involved.

It remains unclear when the observed gender difference develops. The results of several unpublished pilot studies by the present author using students ranging in age from 5 to 13 suggested a possible developmental trend in such differences. Although the sample sizes in these studies were quite small, the older girls' creativity appeared to suffer more when they anticipated evaluation than did the creativity of boys of the same age and under the same conditions. Because of this age-related evidence from the pilot studies, because previous research with middle school students (Baer, 1997) had suggested such a gender difference might exist, and because the period of early adolescence is a very gender-conscious period of development (Gilligan, Lyons, & Hanmer, 1990), middle school (12- to 14-year-old) students were enlisted as subjects in Studies 1-3 to make it most likely that such gender differences would emerge if they did in fact exist.

Second-grade subjects were used in Study 4 because the same pilot studies suggested that gender differences in the impact of anticipated evaluation might not be found among children this young. Although Study 4 could be interpreted as

showing an absence of gender differences among younger (7- to 8-year-old) subjects due to the lack of a gender x motivational condition interaction effect, the fact that boys in Study 4 were *more* creative when they expected evaluation of their work somewhat undermines such a conclusion. This result also suggests the possibility of an additional gender difference in the development of ways that extrinsic motivation influences creativity; that is, it may be that 7- and 8-year-old boys respond to extrinsic constraints in ways that lead to greater creativity, but 7- and 8-year-old girls do *not* respond differently under different motivational constraints. (In contrast, among middle school students it is the girls who appear to respond differently to extrinsic constraints; however, that response results in lower, rather than higher, creative performance.) Because only one study has been done to investigate such gender differences in the effects of extrinsic constraints on primary-school-aged children, however, it would be premature to theorize extensively about such differences. Nonetheless, it is safe to conclude that available evidence suggests that extrinsic constraints generally have a less negative impact on the creative performance of boys than of girls.

There is some evidence suggesting that the observed gender difference may last beyond the period of early adolescence, although this evidence is in the form of differential impact of extrinsic constraints on task motivation rather than creative performance. Kohn (1993), for example, argued that the available evidence points to fairly consistent gender differences in how males and females respond to praise. In terms of the impact of rewards on intrinsic motivation, he claimed that "in general, praise is more likely to have undesirable consequences for females than for males" (Kohn, 1993, p. 106). Deci, Cascio, and Krusell (1975) presented evidence about the differential effects of rewards on males and females that would support the third interpretation above (i.e., that early adolescent girls and boys may *respond differently* to cues from their social environment regarding the appropriate type of motivation). They claimed that, in general, "positive feedback increases the intrinsic motivation of males, whereas it decreases the intrinsic motivation of females" (Deci, Cascio, & Krusell, 1975, p. 84). In two investigations of the effects of praise involving college students (Koestner, Zuckerman, & Koestner, 1987) and upper elementary school students (Koestner, Zuckerman, & Koestner, 1989), similar differences were found. In both of the Koestner et al. studies, praise had a more negative impact

on the females, and it sometimes had a positive effect on the males.

There has been recently a lively debate about the proper use of rewards in education (e.g., Baer & Baer, 1996; Kohn, 1991a, 1991b, 1993; Slavin, 1991a, 1991b). This debate needs to be broadened to consider not only the potential positive and negative effects of rewards in general but also how extrinsic constraints of all kinds should be used with different groups of students. The negative effects of extrinsic motivation include both undermining creative performance and lessening intrinsic motivation (Amabile, 1996). These are serious consequences, and they are the result of conditions that are quite common in the lives of children, especially in school settings. Previous research has identified this problem without reference to possible gender differences, and recommendations have been made of ways to deal with it in the classroom (Baer, 1996; Hennessey & Zbikowski, 1993). The investigations reported above suggest that this problem may be limited to girls, especially among middle school students. Teachers who routinely evaluate the schoolwork of girls of this age and offer rewards for such work need to bear in mind that this may have a significant impact on the creative performance of girls and, quite probably, on their levels of intrinsic motivation as well.

In a typical middle school classroom, it would of course be difficult — and quite possibly counter-productive — to treat boys and girls differently in terms of the kinds of rewards offered and methods of evaluation employed. A wise solution might be to use both rewards and evaluation as little as possible, and, when the use of extrinsic motivators is deemed necessary, to emphasize informational, enabling feedback rather than controlling extrinsic motivators (Amabile, 1996; Deci and Ryan, 1985).

- REFERENCES
- ABELSON, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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). *The social psychology of creativity*. New York: Springer-Verlag.
- AMABILE, T. M. (1985). Motivation and creativity: Effects of motivational orientation on creative writing. *Journal of Personality and Social Psychology*, 48, 393-399.
- AMABILE, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview.

- AMABILE, T. M., GOLDFARB, N., & 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.
- BAER, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal*, 4, 23-39.
- BAER, J. (1993). *Divergent thinking and creativity: A task-specific approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- BAER, J. (1994). Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Research Journal*, 7, 35-46.
- BAER, J. (1996). *Creative teachers, creative students*. Boston: Allyn and Bacon.
- BAER, J. (in press). Gender differences in creativity. In M. A. Runco (Ed.), *Creativity Research Handbook* (Vol. 3). Cresskill, NJ: Hampton Press.
- BAER, J. (1997). Gender differences in the effects of anticipated evaluation on creativity. *Creativity Research Journal*, 10, 25-31.
- BAER, J., & BAER, S. (1996). Are group rewards necessary for cooperative learning? An empirical investigation. In J. A. Chambers (Ed.), *Selected papers from the seventh national conference on college teaching and learning* (pp. 1-6). Jacksonville, FL: Florida Community College.
- BEM, D. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1-62). New York: Academic Press.
- BERGLAS, S., AMABILE, T. M., & HANDEL, M. (1981). *Effects of evaluation on children's artistic creativity*. Unpublished manuscript, Brandeis University.
- CALDER, B., & STAW, B. (1975). Self-perception of intrinsic and extrinsic motivation. *Journal of Personality and Social Psychology*, 31, 599-605.
- CHEEK, J. M., & STAHL, S. Shyness and verbal creativity. *Journal of Research in Personality*, 20, 51-61.
- CONTI, R., & AMABILE, T. M. (1995). *Problem solving among computer science students: The effects of skill, evaluation expectation and personality on solution quality*. Paper presented at the annual meeting of the Eastern Psychological Association, Boston, MA.
- CONTI, R., COLLINS, M. A., & PICARIELLO, M. (1995). *Differential effects of competition on the artistic creativity of girls and boys*. Unpublished manuscript, Brandeis University.
- DECHARMS, R. (1968). *Personal causation*. New York: Academic Press.
- DECI, E. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18, 105-115.
- DECI, E., CASCIO, W. F., & KRUSELL, J. (1975). Cognitive evaluation theory of some comments on the Calder and Staw critique. *Journal of Personality and Social Psychology*, 31, 81-85.
- DECI, E., & RYAN, R. M. (1985). *Intrinsic Motivation and self-determination in human behavior*. New York: Plenum.

- GILLIGAN, C., LYONS, N. P., & HANMER, T. J. (Eds.). (1990). *Making connections: The relational world of adolescent girls at Emma Willard School*. Harvard University Press.
- HENNESSEY, B. A., & AMABILE, T. M. (1988). Conditions of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 11-38). Cambridge University Press.
- HENNESSEY, B. A., AMABILE, T. M., & MARTINAGE, M. (1989). Immunizing children against the negative effects of reward. *Contemporary Educational Psychology*, 14, 212-227.
- HENNESSEY, B. A., & ZBIKOWSKI, S. M. (1993). Immunizing children against the negative effects of reward: A further examination of intrinsic motivation training techniques. *Creativity Research Journal*, 6, 297-307.
- KELLY, H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation*, (Vol. 15, pp. 192-238). Lincoln, NE: University of Nebraska Press.
- KELLY, H. (1973). The processes of casual attribution. *American Psychologist*, 28, 107-128.
- KOESTNER, R., ZUCKERMAN, M., & KOESTNER, J. (1987). Praise, involvement, and intrinsic motivation. *Journal of Personality and Social Psychology*, 53, 383-390.
- KOESTNER, R., ZUCKERMAN, M., & KOESTNER, J. (1989). Attributional focus of praise and children's intrinsic motivation. *Personality and Social Psychology Bulletin*, 15, 61-72.
- KOHN, A. (1991a). Cooperative grade grubbing versus cooperative learning. *Educational Leadership*, 48(5), 83-87.
- KOHN, A. (1991b). Don't spoil the promise of cooperative learning. *Educational Leadership*, 48(5), 93-94.
- KOHN, A. (1993). *Punished by rewards*. Boston: Houghton Mifflin.
- LEPPER, M., & GREENE, D. (1975). Turning work into play: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. *Journal of Personality and Social Psychology*, 31, 479-486.
- LEPPER, M., & GREENE, D. (1978). *The hidden costs of reward*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- LEPPER, M., GREENE, D., & NISBETT, R. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28, 129-137.
- Pool, R. (1994). *Eve's rib: Searching for the biological roots of sex differences*. New York: Crown.
- ROSS, M. (1975). Salience of reward and intrinsic motivation. *Journal of Personality and Social Psychology*, 32, 245-254.
- SLAVIN, R. E. (1991a). Group rewards make groupwork work. *Educational Leadership*, 48(5), 89-91.
- SLAVIN, R. E. (1991b). Synthesis of research on cooperative learning. *Educational Leadership*, 48(5), 71-82.

John Baer, Memorial 102, Rider University, Lawrenceville, NJ 08648 (E-mail: baer@enigma.rider.edu).