

# The Journal of Early Adolescence

<http://jea.sagepub.com>

---

## Long-Term Effects of Creativity Training with Middle School Students

John M. Baer, Jr.

*The Journal of Early Adolescence* 1988; 8; 183


DOI: 10.1177/0272431688082006

The online version of this article can be found at:

<http://jea.sagepub.com/cgi/content/abstract/8/2/183>

---

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *The Journal of Early Adolescence* can be found at:

Email Alerts: <http://jea.sagepub.com/cgi/alerts>

Subscriptions: <http://jea.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

## Long-Term Effects of Creativity Training with Middle School Students

John M. Baer, Jr.  
Rutgers University

### Abstract

The effectiveness of a creativity-training program was investigated. Two matched eighth-grade classes of high-ability students ( $N = 48$ ) were chosen for the study. The experimental group spent three days and two nights at an outdoor school learning to use the Osborne-Parnes model of creative problem solving. There was no treatment for the control group. Just prior to the training both groups were tested, and six months after the training they were retested. There were four subtests: (a) data-finding, (b) problem-finding, (c) idea-finding, and (d) solution-finding, containing both divergent and convergent sections. Mean gain scores of the two groups were compared. On every subtest the experimental group outperformed the control group significantly, indicating that after six months the skills of creative problem solving had been retained.

There is a growing literature on research in creativity and the effects of creativity training. Treffinger (1986), Draper (1985), and Torrance and Presbury (1984) have surveyed current research trends in creativity. Treffinger concluded that the importance of creative productivity is being increasingly recognized, and that “improving students’ creative thinking and problem solving abilities are viable educational goals” (1986, p. 16).

Despite growing recognition of the value of creativity training, evaluations of creativity training programs have been criticized because they have failed to show whether increased creativity scores were attributable to an enduring result of the training in question, or were merely artifacts of just-completed practice sessions on problems much like those tested (Mayer, 1983; Treffinger, 1986; Amabile, 1983). Most studies have concentrated on short-term effects but have failed to assess long-term consequences of creativity training.

Research in creativity has also been criticized because of disagreement about how best to measure creativity (Amabile, 1983; Anderson, 1980; Mayer, 1983; Parnes & Noller, 1973; Parnes & Noller, 1974; Torrance & Presbury, 1984; Treffinger, 1986; Williams, 1980). The most commonly used measures of creativity have been the Torrance Tests of Creativity (Torrance, 1966; Torrance & Presbury, 1984; Amabile, 1983), which assess four divergent-thinking skills: (a) fluency (the production of a large number of ideas), (b) flexibility (the production of a wide variety of ideas), (c) elaboration (the development and embellishment of ideas), and (d) originality (the production of uncommon ideas). Ward (1974), Mayer (1983), and Amabile (1983) have pointed out that tests like the Torrance Tests measure only component abilities that are theorized to contribute to creativity, not creativity per se. Torrance himself has argued for a wide variety of indicators of creativity (Torrance, 1984), and notes a new trend in recent research toward use of more "real-life" measures in studies involving adults (Torrance & Presbury, 1984). Amabile's series of studies exemplified this technique, using experts in a variety of domains to assess the creativity of products such as collages and poems.

A related objection to tests such as the Torrance Tests has been that, while adequately assessing several divergent thinking skills, they have ignored convergent or critical-thinking skills necessary to creative productivity. *Divergent thinking* refers to thinking that moves outward from a problem in many possible directions, such as might be used in answering the question, "How many uses can you think of for a brick?" *Convergent thinking* refers to thinking that proceeds toward a single answer, such as might be used in answering the question, "What is the square root of 25?" (Guilford, 1967). As Treffinger (1986) has noted, "both sets of skills [divergent and convergent thinking] must be used in harmony" (p. 18). Almost all available tests of creativity, however, are "similar in form, content, administration, and scoring" (Amabile, 1983, p.22) to the Torrance Tests; that is, they test only divergent thinking.

The results of research in brainstorming, the most commonly used technique to improve divergent thinking, have been inconclusive (Mayer, 1983), and have highlighted the problem of assessing only divergent thinking skills. In one study (Johnson, Parrott, & Stratten, 1968), subjects using brainstorming produced both a greater number of solutions and more high-quality solutions (as judged by independent raters) than subjects in a control group, indicating better performance in divergent thinking. The subjects in this experiment failed a follow-up test of convergent thinking, however, in a manner that called into question the value of improving divergent thinking without also improving convergent thinking. The test of convergent thinking used by Johnson et al. required the subjects to choose the best solutions from the list of possible solutions the subjects had previously generated. When the subjects who had used brainstorming were asked to choose their own best solutions, the choices were no better than the choices of a control group that chose from its own, shorter list of possible solutions. These results demon-

strated the need to combine the teaching of convergent thinking skills in conjunction with instruction in divergent thinking.

The purpose of the present study was to evaluate the long-term effectiveness for young adolescents of a particular creativity training program. Prior to the training for the experimental group, a pretest of creativity emphasizing both divergent and convergent thinking skills was administered to the experimental group and the control group. Six months later a posttest comparable to the pretest was administered to both groups. It was hypothesized that (a) the experimental group would show greater gain scores than the control group on the tests of creativity, and (b) that the experimental group would show greater gain scores than the control group in each of four subtests of creativity.

## METHOD

### Subjects

The students in both experimental and control groups were eighth-grade students (12-13 years of age) at two similar middle schools, with the determination of experimental and control groups made by the flip of a coin. The middle schools were city schools in a predominately rural county of a Middle Atlantic state, served largely lower-income neighborhoods (in comparison with national, state and county average incomes), and were the only two middle schools of eight middle schools in the county that scored below the national averages on the California Achievement Tests (CAT; Harris, 1978) and Otis-Lennon School Ability Test (Otis & Lennon, 1979).

In each school the "top" section was chosen to participate in the study. Top sections were composed in both schools of those students whose math ability (as assessed by tests and math teachers) qualified them to study algebra in the eighth grade. In the school of the experimental group, the top section included approximately 11 percent of the eighth-grade class. In the school of the control group, the top section included approximately 17 percent of the eighth-grade class. Individual IQ and achievement test scores for students in these two classes were not available, so although the two school populations seem to be well matched, the exact comparability of the experimental and control groups cannot be assured. (The average IQ of all eighth-grade students at both the experimental and control group's schools was 99. The eighth-grade students at the control group school had an average battery total of 8.4 on the CAT; the eighth grade at the experimental group school had an average of 8.3. The national norm for that test was 8.5.)

None of the students in the experimental group reported having taken part in any formal training in creativity or problem solving prior to the seminar. A few of the students had heard of brainstorming, but none could explain what was meant by the term "brainstorming."

In the six-month interval between the pretest and posttest considerable attrition occurred in both groups. The attrition occurred because (a) five students moved to

different school districts, and (b) five students were absent the day of the posttest. The absent students were removed from the study to avoid contamination of the data (as the absent students might then have had prior knowledge of the test items). The number of students in the control group decreased from 25 to 20, and the number in the experimental group dropped from 33 to 28.

### Procedure

The experimental group spent three days and two nights at the county outdoor school, where the seminar was held, and then returned to their regular classes. The control group received no treatment and did not attend the outdoor school. Both groups were tested just before the experimental group received the training, and both were retested six months later.

The model of problem solving taught was *Creative Problem Solving* (CPS) as developed by Parnes and his associates (Parnes, 1972; Biondi, 1972). The training emphasized both divergent and convergent thinking; that is, not only was the production of many *imaginative* ideas encouraged, but the selection and development of *workable* ideas were also stressed. Problems from math, language arts, science, social studies, and the out-of-school lives of students were the content to which students applied the problem-solving techniques they were taught.

For two and one-half days the students worked on a variety of problems, beginning with demonstration problems devised by the instructor and progressing through problems in science, math, language arts, and social studies, with considerable attention given to individual, student-generated problems. The CPS model was introduced through total-class problem solving and reinforced in all problem-solving work done by the students. Individually and in groups students practiced the several steps of the model, sometimes focusing on one step in the process, sometimes using the steps in sequence as part of a complete problem-solving exercise. Each step involved both divergent and convergent thinking phases, and an attempt was made to emphasize the phases equally.

CPS is a five-step model of creative problem solving which includes:

1. *Data-finding*— gathering a wide variety of information, including both hard “facts” and also feelings about a situation, and selecting the most pertinent data and questions; it precedes problem definition so that potentially relevant data isn’t excluded by a narrow or premature definition of the problem.
2. *Problem-finding*— stating a problem in terms of desired outcomes (“In what ways might I...?”) in a way that allows the widest range of possible solutions; generating many different problem statements and then selecting the best.
3. *Idea-finding*— producing many possible solutions to a problem before choosing a solution; the most common technique is brainstorming, but there are many others.
4. *Solution-finding*— developing criteria with which to judge ideas; choosing the criteria that are the most relevant to a problem and applying those criteria to proposed solutions.

5. *Action-planning*—determining (a) a sequence of steps for implementing a solution, (b) whose help will be needed to implement a solution, and (c) who will carry out each part of the solution; establishing a schedule that will make it possible to assess progress in implementing a solution and to evaluate the success of a solution.

The CPS model was the framework for all instruction during the seminar, but the following techniques were also employed:

1. A *Synectics* approach was used in teaching idea-generating techniques for writing. *Synectics* (from the Greek word for joining what appear to be unrelated elements) was developed by Gordon (1961). The approach features a use of a variety of analogies to help idea generation, and like brainstorming encourages deferred judgment during idea generation. Two guiding principles of *Synectics*—making the familiar strange (taking something commonplace and finding new ways of thinking about it through analogy), and making the strange familiar (taking a new problem and transforming it into something familiar by use of analogy)—were emphasized during the seminar.

2. The *scientific method* (Ramsey, Gabriel, McGuirk, Phillips, & Watenpugh, 1986) was compared to the CPS model, and students designed science experiments following the steps of CPS and the scientific method.

3. Students solved a variety of *Minute Mysteries* such as those in *Beyond the Easy Answer* (Weintraub & Krieger, 1978). *Minute Mysteries* were used to sharpen data-finding skills by improving the quality of students' questions. The following is an example of a *Minute Mystery* of the type used during the seminar:

A man is found face down in a field with a pack  
on his back. He is dead. What happened?

Students tried to solve each mystery while asking the fewest possible number of questions. (In the example, the solution is that the man's parachute had failed to open.)

4. Students solved a variety of *math puzzles*, including one made up by a student.

5. Students played a variety of problem-solving games, such as *Charades* (a communication game in which one member of a team acts out nonverbally the title of a book, song, or movie and teammates guess the title).

### Instrumentation

Both pretest and posttest had four parts, corresponding to the first four steps of the CPS model. The tests were paper-and-pencil tests constructed by the instructor. The final step (action-planning) was not included because of difficulties in devising a time-limited, paper-and-pencil test.

The instructions for the pretest and posttest were identical. The specific problems, however, were different. Students' test papers were assigned numbers randomly, so that the two scorers (who were trained together but did their scoring independently) had no indication of which papers came from which school. Inter-

rater reliability for all subtests were in an acceptable range (.72 to .92). Pearson reliability coefficients for each subtest are reported in the following sections.

The four subtest scores were computed by adding the scores of the two raters, with a maximum score on each subtest of 100. The total-battery score was the sum of these four subtest scores.

Gain scores were computed for each student by subtracting the pretest score on each subtest from the posttest score on that subtest. These were used to calculate mean gain scores for the experimental and control groups. Individual and mean gain scores for the total battery of tests were computed in the same way.

The four subtests were timed individually. The rationale for timing the tests was not theoretically based; that is, it was not assumed that being able to answer quickly suggests that one is creative. Rather, the decision to time the tests was a practical one. As the time constraints were uniform for both groups in both testing situations and the key comparisons were between the performance of the same students on pretest and posttest, it seems reasonable to assume that distortions due to time pressures of the key variable (gain scores between pretest and posttest) would have been negligible.

The test instructions and questions are listed below, with brief explanatory notes for each.

#### *1. Data-Finding Test Instructions:*

Below is a problem that you are to help solve. In this test, you are to think of as many questions as you can that you might want to have answered before solving the problem. At the end of four minutes you will be told to begin selecting the three most important questions. You will then have two minutes to circle those three.

Pretest: There has been a big increase in the number of students dropping out of school.

Posttest: You are the principle of an elementary school. One of your students brought a pet snake to school today, and just discovered that it is missing from its cage.

An examination of student responses to the pretest question highlighted a problem of interpretation (a threat to validity that was corrected in the posttest). Some students looked at the problem from the perspective of an eighth-grade student, others from that of an educator or parent, and a few took the question very personally, as if they were deciding whether or not to drop out of school themselves.

The interrater reliability of both tests was acceptable (.86 for the pretest, .86 for the posttest).

*Data-finding test scoring:* (a) Quantity (0-25): Count the number of different questions. Maximum score is 25; (b) Variety (0-10): Do the questions cover a wide range, including relevant facts, background information, important people, goals?; (c) Originality (0-5): Are there questions on the list that are unusual?; (d) Selectivity (0-15): How helpful would the three circled questions be in assessing this situation?

## 2. Problem-Finding Test Instructions:

The next test is to find out how many different ways you can think of to state the problem. Start each question with "How can we" or "How can I" and then write the problem. You should try to find as many different ways to state the problem as you can. After four minutes you will be given two more minutes to select and circle the best problem statement. Use the same problem that you used in the last (data-finding) test.

Additional verbal instructions were given at each testing situation as follows: "For example, if the problem were mice in my basement, I might ask 'How can I build a better mousetrap?' 'How can we get rid of the mice?' or 'How can I not be bothered by the mice?'" This was the result of questions asked by subjects the first time the test was given (to the experimental group). To be consistent, the same instructions were repeated each time.

Reliability coefficients were acceptable (.72 for the pretest and .91 on the posttest).

*Problem-Finding test scoring:* (a) Quantity (0-25): Count the number of different problem statements and multiply by two. Maximum score is 25; (b) Originality (0-10): Are there unusual or interesting problem statements?; (c) Selectivity (0-15): How good is the problem statement selected? Is it broad enough to allow for many possible solutions? Does it include most relevant aspects of the problem?

## 3. Idea-Finding Test Instructions

You will be given a problem and asked to think of as many different answers as you can. You have five minutes.

Pretest: How might we stop wars?

Posttest: How could schools be improved?

As each student selected his own problem statement in the problem-finding test (and thus these problems would vary in difficulty), the problem was changed for this and the next (solution-finding) test. This insured that each testing situation was the same for all students.

No convergent-thinking criterion was used for this test.

Due to the simple scoring system for this test, the two raters' scores were nearly identical. No correlation coefficient was calculated.

*Idea-Finding test scoring:* Count the number of different answers and multiply by two. Maximum score is 50.

## 4. Solution-Finding Test:

List things you might consider in deciding which of your ideas in the last (idea-finding) test are best. What standard will you use to judge them? After five minutes you will be asked to circle the three that you think are the most important. You will have two minutes to do this.

As in the problem-finding tests, additional explanation was called for at the first test site, and this was repeated at each testing. The example given was: "If you were going to buy a pair of shoes, there are a number of things you would want to

consider. 'Do they fit?' 'Can I afford them?' 'Are they a style I like?' 'Will they last?' 'Will my parents approve?' Different people would use different criteria, of course. You are to list all the things you might consider in deciding which of your solutions to this problem are best. Write as many as you can think of.

Reliability coefficients were acceptable: .82 for the pretest and .92 for the posttest.

*Solution-Finding test scoring:* (a) Quantity (0-30): Count the number of different criteria and multiply by two. Maximum score is 30; (b) Appropriateness (0-20): How appropriate to this problem are the three criteria selected? How well do they cover the range of the factors which need to be considered in picking a solution?

This study was the first use of these test, and there is therefore no normative data. The control group scored only slightly higher on the posttest than they did on the pretest, however, indicating that the degree of difficulty was probably not greatly different. As the only comparisons made in this study were between average increases in score from the first test to the second, it was not important that they be of precisely equal difficulty.

It must be noted that this study was not a test of the CPS model. Rather, CPS was assumed to be a reasonable description of the problem-solving process, and the effectiveness of the seminar was evaluated using the components of the model. Two aspects of the study are of special interest in that (a) three of the four subtests used in the study combined divergent-thinking skills (which are typically the only skills assessed in creativity studies) with convergent (critical) thinking to provide a more full assessment of subjects' creative problem-solving abilities, and (b) the posttest was given not immediately following the seminar but six months later.

## RESULTS

The experimental group, in testing six months after the seminar, showed significantly higher increases on each of the subtests and the total battery, using a one-tailed *t*-test.

The control group showed a small overall gain of 5.95 points. On the 400-point total battery, this amounts to approximately 1.5 percent of the total possible score. The experimental group had a mean gain score of 76.29, or 19.1 percent of the maximum score (See Table 1).

A difference score was computed for each subtest by subtracting the control group's mean gain score from the experimental group's mean gain score. The difference scores for the four subtests were 26.95 (Data-Finding), 17.23 (Problem-Finding), 14.14 (Idea-Finding), and 15.01 (Solution-Finding). These difference scores can also be read as percentages (of the subtest maximum score of 100) to give an indication of the magnitude of the differences between experimental and control group performance. Similarly, the maximum possible total-battery score was

TABLE 1  
Mean Gain Scores of Experimental and Control Groups

Group	Data-Finding	Problem-Finding	Idea-Finding	Solution Finding	Total Battery
Experimental ( <i>n</i> = 28)	13.500	23.929	24.143	14.714	76.286
Control ( <i>n</i> = 20)	-10.450	6.700	10.000	-0.300	5.950
Difference (One-tailed <i>t</i> -test)	23.950 <i>p</i> <.0005	17.229 <i>p</i> <.0005	14.143 <i>p</i> <.005	15.014 <i>p</i> <.005	70.336 <i>p</i> <.0001

400 and the difference score between groups was 70.34, which was 17.6 percent of the maximum possible score. The significance levels of the difference scores were all  $p < .005$  ( $p < .0001$  for the total battery).

## DISCUSSION

The increase in the problem-solving skills of the experimental group was both substantial and statistically significant. In fact, every student in the experimental group scored a larger total-battery gain score than the average total-battery gain score of the control group.

No measure of attitudes was taken, however, both the seminar instructor and the two classroom teachers who observed the seminar stated that the students seemed to have enjoyed the learning experience. The observer-teachers also expressed the belief at the time of the posttest that the students who had been in the seminar class were more creative (and, interestingly, more cooperative) than students in classes of similar ability which they had taught in the past.

The two observer-teachers expressed great interest in creativity training at the end of the seminar, however, in a follow-up interview at the time of the posttest the two teachers agreed that plans to integrate CPS into classroom teaching had been largely forgotten shortly after returning to the regular school classes. It is certainly possible that some of the techniques that had been observed had unconsciously become part of the classroom teaching; in fact, the two teachers suggested that this probably had occurred. The degree to which the results can be attributed to the students' experiences at the two-and-one-half-day creativity seminar is not clear, therefore, as a change in the teaching practices of the two observing teachers may have been an important factor. To assess the relative impacts of the seminar itself and classroom follow-up by observer-teachers and to determine the importance of having classroom teachers observe such a seminar, this study might be replicated with a group of students whose teachers did not observe the seminar. The control

group in such a replication should spend a like amount of time at the outdoor school involved in activities not related to creativity training to reduce the possibility of a Hawthorne effect for the experimental group.

The results of the current study suggest that creativity training might prove to be a valuable adjunct to the regular school program, and reports of observers support this conclusion. Similar seminars had been conducted with six groups ranging from fourth through eighth grades prior to this study. Teachers had been unanimous in their approval of the training, and had reported that students who had learned CPS methods had shown impressive gains in problem-solving skills needed in school, as well as improvement in interpersonal relationships. Future studies may clarify (a) which elements of the training should be emphasized, (b) the effects of CPS follow-up in the classroom, (c) the ages best suited to CPS training, (d) the degree to which the benefits of CPS training can be generalized to average and below-average achievers, and (e) the effectiveness (including cost-effectiveness) of offering CPS training in the regular school rather than isolating students at the outdoor school.

## REFERENCES

- Amabile, T. M. (1983). *The social psychology of creativity*. New York: Springer-Verlag.
- Anderson, John R. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.
- Biondi, A. M. (Ed.) (1972). *The creative process*. Buffalo, NY: D. O. K.
- Draper, W. D. (1985). A cursory review of current creativity research. *Creative Child & Adult Quarterly*, 10, 86-89.
- Gordon, W. J. J. (1961). *Synectics*. New York: Harper & Row.
- Guilford, J. P. (1967) *The nature of human intelligence*. New York: McGraw-Hill.
- Harris, J. (Senior Project Director) (1978). *California Achievement Test, Form C*. Monterey, CA: McGraw-Hill.
- Johnson, D. M., Parrott, G. L., & Stratten, R. P. (1968). Production and judgment of solutions to five problems. *Journal of Education Psychology*, 59, Monograph Supplement No. 6.
- Mayer, Richard E. (1983). *Thinking, problem solving, cognition*. New York: Freeman.
- Otis, A. & Lennon, R. (1979). *Otis-Lennon School Ability Test*. San Antonio: Psychological Corporation.
- Parnes, S. J. (1972). *Creativity: Unlocking human potential*. Buffalo, NY: D. O. K.
- Parnes, S. J. & Noller, R. B. (1973). *Toward supersanity*. Buffalo, NY: D. O. K.
- Parnes, S. J. & Noller, R. B. (1974). *Toward supersanity: Research supplement*. Buffalo, NY: D. O. K.
- Ramsey, W. L., Gabriel, L. A., McGuirk, J. F., Phillips, C. R., & Watenpaugh, F. M. (1986). *Physical science*. New York: Holt, Rinehart & Winston.
- Torrance, E. P. (1966). *The Torrance tests of creative thinking: Norms—technical manual*. Lexington, MA: Personal Press.
- Torrance, E. P. (1984). The role of creativity in identification of the gifted and talented. *Gifted Child Quarterly*, 28, 153-156.
- Torrance, E. P. & Presbury, J. (1984) The criteria of success used in 242 recent experimental studies of creativity. *Creative Child & Adult Quarterly*, 9, 238-243.
- Treffinger, D. J. (1986) Research on creativity. *Gifted Child Quarterly*, 30, 15-19.

- Ward, W. C. (1974). Creativity (?) in young children. *Journal of Creative Behavior* 8, 101-106.  
Weintraub, R. & Krieger, R. (1978). *Beyond the easy answer*. Washington, DC: Zenger.  
Williams, F. (1980). *Creativity assessment packet manual*. Buffalo, NY: D. O. K.

Reprint requests should be addressed to:

John Baer  
Department of Psychology  
Rutgers University  
New Brunswick, NJ 08903