selected in the previous stage are prepared in a plan for action.

The Type III management plan requires students to prepare a list of what they will need to get started, resources they will consult and what action they will take to develop their products and performances. Goals and timelines are included in this.

Working through an acceptance-finding stage in anticipation of starting their research enabled the environmental group to list ways to assure the drafted project would succeed. Initial obstacles were identified through this last stage of CPS: scheduling time to meet, designating responsibilities, and locating mentors and equipment All of these were listed on their management plan and became the first tasks at hand.

Implications

Research has shown that students who are given specific training in the methodology of a Type III investigation (the mechanics of what it involves) have an increased likelihood of participation in the process of creating original products in response to investigations of real world problems (Burns, 1987). In a recent study, students of varying abilities were shown to have increased problem-solving ability from instruction in Creative Problem Solving and its application to a real school problem (Schack, 1993). One implication of this article is that research needs to be conducted to determine whether instruction in the stages of CPS along with strategies that relate it to Type III production and integration of the two processes does enhance the frequency and quality of student creative production.

It has been my experience as a teacher that use of the integrated approach discussed here has increased the number of advanced level research projects I have facilitated in both the regular classroom and resource room. Through use of the CPS heuristic, the problem-finding—problem-solving nature of creative performance or product development has been underscored for the learners I have worked with. Guiding students with CPS to generate ideas for Type IIIs and evaluate them by criteria related to authentic problem solving has, in my opinion, enhanced student creative productivity. These experiences support my strong recommendation that the two models be integrated.

REFERENCES

Cray-Andrews, M., Baum, S. E & Gubbins, J. E. (1987). Gifted students as creative producers, *GCT*. (pp. 22-25).

Baron, J. B. & Sternberg, R. J. (1987). Introduction. In J. B. Baron and R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice*. (pp.1-5).New York: W. H. Freeman.

Baron, J. B. & Sternberg, R. J. (1987). Questions and answers about the nature and teaching of thinking skills. In J. B. Baron and R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp.251-261). New York: W. H Freeman.

Beyer, B. K. (1991). *Teaching thinking skills: A handbook for elementary school teachers.* Boston: Allyn and Bacon.

Burns, D. E. (1990). *Pathways To Investigative Skills.* Mansfield, Connecticut: Creative Learning Press.

Cramond, B., (1990). Generalizability of creative problem solving procedures to real- life problems., *Journal for the Education of the Gifted, 13*, pp. 141-155.

Cramond, B., Martin, C., & Shaw, E. L. (1988). An investigation of the application of training in creative problem solving to content area problems. In *American Educational Research Association, Annual Meeting* [Microform]. New Orleans, LA.

Feldhusen, J. F. & Treffinger, D. J. (1977). *Creative thinking and problem solving and gifted education.* Dubuque, Iowa: Kendall/Hunt.

Isaksen, S. G.& Treffinger, D. J. (1985). *Creative problem solving: the basic course.* Buffalo: Bearly Limited.

Maker, J. C. (1982). *Teaching models in gifted education.* London: Aspen Publications.

Maker, J. C. (1986). Integrating content and process in gifted education. In J. C. Maker (Ed.), *Critical Issues in Gifted Education* (pp. 151-163). Rockville: Aspen Publications.

Renzulli, J.S. , (1982). *What makes a real problem: Stalking the illusive meaning of qualitative differences in gifted education.* Mansfield, Connecticut: Creative Learning Press.

Renzulli, J.S.,(1985). *The three ring conception of giftedness: A developmental model for creative productivity.* Mansfield, Connecticut: Creative Learning Press.

Renzulli, J.S. and Reis, S.M. (1985) *The schoolwide enrichment model.* Mansfield, Connecticut: Creative Learning Press.

Torrence, E. P. (1987). Part One. Teaching for creativity: Can we teach children to think creatively? In S. G. Isaksen (Ed.) *Frontiers of Creativity Research: Beyond the Basics* (pp. 189-204). Buffalo: Bearly Limited.

Torrence, E. P. (1987). Part Two. Teaching for creativity: Recent trends in teaching children and adults to think creatively In S. G. Isaksen (Ed.) *Frontiers of Creativity Research: Beyond the Basics* (pp. 189-204). Buffalo: Bearly Limited.

Performance Assessments of Creativity: *Do They Have Long-Term Stability?*

John Baer

Assessment by experts of the creativity of products (such as poems and collages) has recently become an important technique for measuring creativity. While inter-rater reliabilities have been high, there have been no studies of the long-term stability of measures of creativity based on the assessment of such products. Two studies are reported that measure the long-term stability of performance assessments involving story-writing and poetry-writing (fourth- and fifth-grade students) and story-telling (second-grade students). The long-term stability of these assessments compares favorably with stability figures for other creativity tests.

John Baer is assistant professor, School of Education and Human Services, Rider College, New Jersey and author of *Divergent Thinking and Creativity: A Task-Specific Approach* (Erlbaum, 1993).

Many new approaches to measuring creativity have appeared in the past decade, of which the most promising involve assessment by experts of the creativity of actual products, such as poems, stories, or collages. Amabile and her colleagues (Amabile, 1982, 1983; Hennessey & Amabile, 1988a, 1988b) have developed and tested these techniques extensively. Over 30 experimental studies have been conducted using such "consensual" techniques — consensual because the evaluations of creativity are based on a consensus of experts in whatever domain is being tapped by the tests.

Hennessey and Amabile (1988b) recommend story-telling as a method for assessing children's creativity in psychological and educational research. Educators have recently begun using similar "real world" measures of creative performance (e.g., Connecticut State Department of Education, 1988), and even Torrance has advocated the use of performance measures

Manuscript submitted April, 1993.
Revision accepted October, 1993.

as one measure of creativity (Torrance & Presbury, 1984). Consensual assessments of creative products will not fulfill all creativity-assessment needs, but they may help us avoid the "Creativity Quotient Fallacy" (Treffinger, 1986, p. 16): the belief that a single test can tell us all we need to know about a person's creative potential.

Consensual assessment of creative products has many advantages over the most common technique for creativity measurement, divergent-thinking tests like the Torrance Tests of Creative Thinking (Torrance, 1966; Torrance & Presbury, 1984). These tests are based on a popular theory of creativity (Guilford, 1967; Guilford & Hoepfner, 1971) that stresses the importance of divergent thinking in creativity. The relationship between divergent thinking and creativity is not a simple one, however (Kogan, 1983), and there are many competing theories of creativity that deny divergent thinking a primary role in creative thought. Tardif and Sternberg (1988) summarize several theories of creativity that are not based on Guilford's divergent-thinking model, and Baer (1993) reviews research that casts doubt on the validity of divergent-thinking tests as measures of creativity.

There is also a growing body of evidence that creativity is not a generic trait or set of skills that apply regardless of the task domain. Recent research suggests that creative-thinking skills are often very narrowly task-specific (Baer, 1991, 1992, 1993; Runco, 1986). This research calls into question the validity of generic creative-thinking tests (such as most divergent-thinking measures). It also points to the need for educators responsible for selecting students for gifted/talented programs to tailor the kinds of creativity testing they employ to the specific kinds of programs they are offering.

Unlike divergent-thinking tests of creativity, consensual assessments of creative products are not grounded in any theory of creative thinking; the only "theory" upon which they are based is the belief that experts in a given field can recognize creativity when they see it. This is, in fact, how creativity is typically assessed in almost all fields, even the "hard" sciences (Kuhn, 1970). Because experts' assessments of creativity are independent of the creativity theories of psychologists, the validity of these assessments is not contingent on the validity of any theory of creativity. In fact, divergent thinking might be shown to have no important or meaningful relationship to creativity at all, and the validity of consensual approaches to creativity assessment would remain the same.

The procedure is fairly simple: subjects are asked to produce something — a poem, story, or collage, for example — and these are rated by experts for creativity. High inter-rater reliabilities have been the norm, and creativity scores obtained in this way have been shown to be distinct from ratings of technical goodness, factual detail, complexity, or other dimensions upon which these products can be rated.

There is one substantial drawback to consensual assessment techniques: creativity scores can be compared only within a particular sample. Because scores are based on the ratings of a particular group of experts, who have rated the products in comparison with the other products in the sample, no norms can be established for comparisons with products from other samples. For most research and educational uses, however, within-group comparisons are all that is required.

Consensual assessment techniques have been shown to have high inter-rater reliability when used with poems, stories, and collages (Amabile, 1982, 1983; Baer, 1991, 1992, 1993; Hennessey & Amabile, 1988a, 1988b), and it is reasonable to expect that this will be the case with other kinds of creative products. These techniques initially were developed by Amabile for use in social psychological research. They have been very successful in showing the effects of experimental manipulations on creative performance; however, this susceptibility to changes in situational constraints might make such measures highly unstable predictors of creative performance. Testing this stability was the goal of the two studies reported here.

It should be noted that other tests of creativity are also susceptible to changes in situational constraints. There is, for example, considerable dispute regarding differences in performance between divergent-thinking tests given in timed, test-like formats or in untimed, game-like formats (Hattie, 1980; Kogan, 1983). It has also been shown that Torrance Test scores are highly susceptible to training or practice effects (Lissitz & Willhoft, 1985). This makes such tests fairly easy to fake, given some knowledge of how they are scored. This is a short-coming divergent-thinking tests share with some personality measures of creativity (Ironson & Davis, 1979; Tomlinson & Wilson, 1973).

Creativity measures based on product evaluations by experts, on the other hand, even if susceptible to situational (especially motivational) constraints, are at least immune to faking. Practice effects should also be minimal, because activities like writing poems, telling and writing stories, or creating collages — activities that are frequently "practiced" by many children and adults for their intrinsic pleasure — are different in each situation. Simply learning how the "test" is scored, which can dramatically influence divergent-thinking test scores (Lissitz & Willhoft, 1985), cannot be a factor; otherwise, artists, poets, and writers would long ago have stolen the "scoring key" in their respective fields.

The two studies reported here were undertaken to assess the long-term stability of three consensual techniques for evaluating creativity. In the first study, 19 fourth-grade students were given topics and asked to write poems and stories, which were evaluated for their creativity by experts. The following year the same students were given different topics for poems and stories. These fifth-grade products were then evaluated by a different group of experts. These scores were then compared to the creativity scores of the previous year's poems and stories. In the second study, 38 second-grade students were given a picture book and asked to tell a story to go with it. These stories were tape recorded, transcribed, and then rated by experts for their creativity. Two different picture books, both offering wide latitude in story interpretation, were used, one in October, the other in June. Ratings of the two stories told by each child were then compared.

STUDY 1
Method

Subjects

Ten girls and nine boys were the participants. They were students in the same fourth-grade class in a small suburban public school in the mid-Atlantic region. They were the only fourth-grade class in the school, and were not, therefore, a specially selected population; however, the students were well above average in intelligence, with a mean Otis-Lennon School Ability Index of 120, and an IQ range of 100 to 146. Because there were no significant correlations between sex and any of the four poetry- and story-writing creativity test scores, separate results will not be reported for boys and girls.

Tests

None of the tests were timed. Most students finished both tasks (the poem and the story) in about one hour on both testing occasions. The tests were given in November of the students' fourth-grade year and 11 months later, in October of the students' fifth-grade year.

Poetry-writing tests: Subjects were asked to write an original poem on the topic of "The Four Seasons" (fourth-grade testing, referred to below as "Poetry 1") and another with the topic "The Wind" (fifth-grade testing, "Poetry 2"). The form, style, and length of the poem were not specified. Subjects were told that except for the topic, everything else about the poem was up to them.

Story-writing tests: Subjects were given a drawing as the subject of each story. In the fourth-grade testing ("Story 1"), the drawing depicted two men, one neatly and one casually dressed, approaching the corner of a building from opposite directions. In the fifth-grade testing ("Story 2"), the drawing depicted a man and a woman, both neatly dressed, sitting at a table talking. The students were asked to write an original story in which the figures in the drawings played some part.

Students' poems and stories were rated in each case by five independent raters, who in all cases did not know the students who wrote the poems and stories. The raters included teacher-specialists in gifted education, poets, short-story writers, and college professors of English. Amabile (1982, 1983) has shown that all of these different kinds of experts — roughly classifiable as practitioners in a field, critics and educators — tend to agree on ratings of creativity. In any domain (such as poetry or short fiction), it is in the final analysis the judgments of these kinds of people — practitioners, critics and educators — who determine what is creative, whether the works being judged are student works or masterpieces that redefine what Kuhn (1970) has termed the accepted "paradigm" of a field. In the case of paradigm-shifting creative efforts, there is always the danger that only critics of the future will recognize works of genius, but with student work (such as that used in these studies) there is little danger of this (Baer, 1993).

Different raters were used for the fourth- and fifth-grade tests. This procedure ensured that the stability coefficients would not be artifacts of the particular stylistic tastes of a single set of raters. The raters were informed of the instructions given to the students, and were told that the students were fourth- or fifth-grade elementary-school students. The raters knew that the experiment involved creativity in some way, but did not know the purpose of the experiment. The definition of creativity was expressly left up to the individual raters, who worked independently and had no opportunity to discuss the project with one another prior to making their ratings. All raters were paid for their work.

Each group of 19 papers was ranked for creativity by each rater from 1 to 19, and these rank-ordered scores were then converted using an area-under-the-curve table into normally distributed standard scores, based on the assumption that the creativity of the papers would be normally distributed. Inter-rater reliabilities ranged from .79 to .88 (see Table 1).

Each student's score on each test was the sum of the scores of the five raters for that paper.

Inter-rater Reliabilities

Test	Reliability Coefficient[a]
Poetry 1	.79
Poetry 2	.84
Story 1	.88
Story 2	.85

Note: There were five raters for each test.
[a]Coefficient alpha

Table 1

Procedure

Students were tested in their regular classroom. Students were told that the experimenter, who was not known by the students, was interested in the creative thinking of students like themselves, and that the activities they were going to do would help him understand this better. All were cooperative and appeared to work diligently on all tasks.

The tests were administered to the entire class at once. To avoid confusion, especially at the first testing, the two tests were not both described in detail at the outset. The students were told that there would be two activities, and that after they had worked on the first activity for a while, the second one would be introduced. This procedure made it less likely that the students would conflate the two tasks or focus on one task to the exclusion of the other. The first test (Story 1) was introduced, and after about 20 minutes the students were interrupted to be told of the other test (Poetry 1). The procedure was the same in the second testing, except for reversing the order of the tests. Students were told that they could work on each test as long as they wished, and that they could switch back and forth between the tests as often as they wished.

Results

Table 2 shows the stability of the poetry-writing and story-writing creativity scores after 11 months.

Stability of poetry-writing and story-writing creativity test scores

Poetry 1 - 2	.44	($p < .06$)
Story 1 - 2	.58	($p < .01$)

Table 1

Story-writing creativity scores were somewhat more stable than poetry-writing creativity scores, and the correlation between the two story-writing creativity scores was statistically significant at the .01 level. The correlation between the two poetry-writing creativity scores just missed the .05 level of statistical significance.

Discussion of Study 1

Both tests showed moderate stability as evidenced by their ability to predict scores on similar tests taken 11 months later. Calling this "long-term stability" may raise some eyebrows, as the term was only 11 months; however, with students this young (9-10 years old), 11 months represents a period of considerable growth and change, and these results suggest that these tests assess fairly stable sets of abilities. By way of comparison, long-term stability figures for divergent-thinking tests with older students range from the high .30s to the low .50s over periods of three to five years. IQ test scores, on the other hand, show greater stability, generally ranging from .60 to .80 (Kogan, 1983).

One might question the basis for the stability of the poetry-writing and story-writing creativity scores by questioning the amount of variance in these scores accounted for by IQ. If this were high, the stable component of the poetry-writing and story-writing creativity tests might be largely due to the effects of IQ (the problem of multicollinearity). High correlations between IQ and the poetry- and story-writing creativity tests would also suggest that what was being assessed was not creativity, but intelligence.

Correlations between IQ scores, on one hand, and the poetry-writing and story-writing creativity scores, on the other, were computed. Poetry 1 and Poetry 2 were combined, and Story 1 and Story 2 were similarly summed. Poetry-writing creativity scores were correlated .38 ($p > .10$) with IQ test scores. Story-writing creativity scores were correlated -.19 ($p > .10$) with IQ test scores. Correlations with Reading Achievement on the California Achievement Tests showed a similar pattern of a positive correlation with the poetry-writing creativity scores and a negative correlation with story-writing creativity scores. It does not appear that the stability of IQ test scores caused the stability between the two consensual measures of creativity — especially the more stable one, story-writing.

STUDY 2
Method

Subjects

Thirty-eight second-grade students — the entire second grade of a small suburban elementary school in southern New Jersey — served as subjects. California Achievement Test scores were above average in both reading and mathematics, although there was a wide range of achievement that roughly approximated a normal distribution of test scores. One student came to the school just prior to the time when these tests were given and had therefore not taken the same battery of standardized achievement tests, but of the 37 students who had taken the California Achievement Tests, the distribution of scores was as follows: ***reading achievement*** — 15 in the top quartile, 13 in the second quartile, 8 in the third quartile, 1 in the lowest quartile; ***math achievement*** — 8 in the top quartile, 17 in the second quartile, 11 in the third quartile, and 1 in the lowest quartile; ***language arts skills*** — 9 in the top quartile, 19 in the second quartile, 6 in the third quartile, and 3 in the lowest quartile.

Test

In the first (October) testing, subjects were shown a picture book, the same one used by Amabile (1983) in the studies she conducted to validate the procedure: *A Boy, a Dog, and a Frog* (Mayer, 1967). After looking through the book at their own pace to become familiar with the story, the students were instructed to "tell the story in your own words by saying one thing about each page" while looking at the book's pictures. These stories were later transcribed by the experimenter (except for two that were lost due to garbled speech in one case and tape recorder malfunction in the other) and given to experts to rate for creativity. The experimenter attempted to transcribe the verbal expressions and emphases that the students used in telling the stories as accurately as possible. On a few occasions this was difficult because part of the flavor of the story was conveyed in the tone of voice used by the students, but for the most part this transcription was a straightforward affair. In the June testing, a similar picture book was used (Mayer & Mayer's 1971 *A Boy, a Dog, a Frog, and a Friend*). The same procedure was followed.

The rating procedure was different than the rank ordering used in Study 1. Judges simply rated each collage or story on a 1.00-to-5.00 scale. Other studies (Baer, 1993) had shown that this simpler procedure yielded the same results as the rank-order procedure used in Study 1. There were four raters for the first story-telling test and five for the second. All were teacher-specialists in gifted education, and all had several years experience working with primary-grade students. As in Study 1, the raters did not know the subjects or the specific purpose of the study, and the raters worked independently of one another. All raters were paid.

The coefficient alpha inter-rater reliabilities were .90 for the first testing and .85 for the second testing.

Procedure

Each student was tested individually. Prior to testing the experimenter met with the students in two class groups and explained what they would be doing.

The students had little trouble understanding the tasks. Some students were shy before starting to tell the story. When this happened, they were prompted to "just say one thing about each page."

Results and Discussion of Study 2

There were a total of 36 pairs of stories, because two stories could not be transcribed (as explained above). The correlation between the two scores was .49 ($p < .005$). To test for the effects of multicollinearity, partial correlations were computed after all variance attributable to standardized reading and math achievement test scores had been removed. This reduced the correlation slightly, to .44 ($p < .01$).

As in Study 1, there was a fairly high degree of stability in test scores over this eight-month period — a period of time representing almost one-tenth of the age of the students.

General Discussion

The long-term stability coefficients of the three performance assessments used in Studies 1 and 2 were at least as good as those of commonly used divergent-thinking tests. The correlations noted above — .44 for the poetry-writing test, .58 for the story-writing test, and .49 for the story-telling test — are the raw correlations, without correction for attenuation. A correction for attenuation can be used to estimate the extent to which observed correlations are attenuated by measurement error (Cohen & Cohen, 1983; Nunnally, 1978). To the extent that measurements are unreliable, correlations between those measures will be lessened, and an estimate can be made of what the correlation would have been if perfectly reliable measures had been used. With correction for attenuation, the long-term stability coefficients increase to .66 for the poetry-writing test, .78 for the story-writing test, and .64 for the story-telling test.

There is some controversy about when this correction for attenuation should be applied (Cohen & Cohen, 1983; Nunnally, 1978), and there are especially strong reasons to

doubt that it is appropriate when reporting long-term stability of test scores. It is often used, however, with the effect of increasing the size of these reported coefficients. But even with correction for attenuation, Kogan reports in his survey of the literature that the long-term stability coefficients of divergent-thinking test performance rarely exceed .50 (Kogan, 1983). The long-term stability of the performance measures of creativity used in Studies 1 and 2 compares quite favorably, then, with the long-term stability of divergent-thinking measures of creativity.

Amabile's (1983) research demonstrates clearly that, under different conditions (especially those influencing intrinsic and extrinsic motivation), creativity test scores will vary considerably. This is an important constraint on any measure of creativity, and should be born in mind when giving creativity tests to aid in the selection of students for gifted/talented classes. The results of Studies 1 and 2 suggest that, if given under similar conditions, the poetry-writing, story-writing and story-telling tasks provide reliable and stable measures of creativity. Consensual assessment techniques thus appear to have cleared another hurdle in their acceptance as valuable measures of creativity. And because of the increasing doubts about the validity of divergent-thinking tests as measures of creative potential, educators responsible for assessing students' creativity should consider either replacing, or at least supplementing, divergent-thinking tests with performance assessments of creativity.

Note: This material is based, in part, on work supported under a National Science Foundation Graduate Fellowship, a Rutgers Excellence Fellowship, a Rutgers Graduate Scholars Award, and a grant from the Educational Testing Service.

REFERENCES

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43,* 97-1013.

Amabile, T. M. (1983). *The social psychology of creativity.* New York: Springer-Verlag.

Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal, 4,* 23-39.

Baer, J. (1992, August). *Divergent thinking is not a general trait: A multi-domain training experiment.* Paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Baer, J. (1993). *Divergent thinking and creativity: A task-specific approach* . Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (Second Ed.). Hillsdale, NJ: Erlbaum.

Connecticut State Department of Education. (1988). *Connecticut Assessment of Educational Progress.* Hartford, CT: Author.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence.* New York: McGraw-Hill.

Hattie, J. (1980). Should creativity tests be administered under testlike conditions? An empirical study of three alternative conditions. *Journal of Educational Psychology, 72,* 87-98.

Hennessey, B. A., & Amabile, T. M. (1988a). Conditions of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 11-38). Cambridge University Press.

Hennessey, B. A., & Amabile, T. M. (1988b) Story-telling: A method for assessing children's creativity. *Journal of Creative Behavior, 22,* 235-246.

Ironson, G. H., & Davis, G.A. (1979). Faking high and low scores on the Adjective Check List. *Journal of Creative Behavior, 13,* 139-145.

Kogan, N. (1983) Stylistic variation in childhood and adolescence: Creativity, metaphor, and cognitive styles. In P. H. Mussen (Ed.), *Handbook of child psychology: Vol 3. Cognitive development* (4th ed., pp. 628-706). New York: John Wiley & Sons.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

Lissitz, R. W., & Willhoft, J. L. (1985). A methodological study of the Torrance Tests of Creativity. *Journal of Educational Measurement, 22,* 1-11.

Mayer, M. (1967). *A boy, a dog, and a frog.* New York: Dial Books.

Mayer, M., & Mayer, M. (1971). *A boy, a dog, a frog, and a friend.* New York: Dial Books.

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.) New York: McGraw- Hill.

Runco, M. A. (1987). The generality of creative performance in gifted and nongifted children. *Gifted Child Quarterly, 31,* 121-125.

Tardif, T. Z., & Sternberg, R. J. (1988). What do we know about creativity? In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 429-440). Cambridge University Press.

Tomlinson, R. F., & Wilson, D. L. (1973). The effects of instructions to appear creative on test scores of creativity. *Newsletter for Research in Mental Health and Behavioral Sciences, 15*(2), 29-31.

Torrance, E. P. (1966). *The Torrance tests of creative thinking: Norms-technical manual.* Lexington, MA: Personal Press.

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.

Creativity as a Characteristic of Giftedness: *Teachers See It, Then They Don't*[1]

Scott L. Hunsaker

Scott L. Hunsaker is assistant professor, Gifted/Creative Education, University of Georgia and serves as the Chair of the Creative Division for NAGC.

In this study, teachers saw giftedness as greatly varied, but having the common characteristic of creativity. However, when observing for nomination to gifted programs, teachers focused on classroom performance to a greater degree than creativity. Thus their observations more closely matched the official definition rather than their personal conceptions. The teachers had not actively worked to align the official conception more closely with their own. This inaction was attributed to lack of experience with gifted education, avoiding harm to children, and being patient with change.

[1] Research for this article was supported in part under the Javits Act Program (Grant No. R206R00001) as administered by the Office of Educational Research and Improvement, U.S. Department of Education. Grantees undertaking projects are encouraged to express freely their professional judgements. This article, therefore, does not necessarily represent positions or policies of the government, and no official endorsement should be inferred.

Using teachers as nominators for gifted identification has been controversial for some time (see, for example, Jacobs, 1971; Pegnato and Birch, 1959). On the one hand, much of the research has shown teachers to be a suppressive influence on the identification process. Forum (1980) and Tuttle, Becker, and Sousa (1988) reported that teachers often resent the gifted program and the time taken by the identification process. Such resentment could affect accuracy of teacher judgments. Further Forum (1980) noted that when teachers viewed gifted programs as elitist and intellectualist, they would hesitate to recommend students to the programs for fear the programs would exacerbate the social problems they felt gifted children have.

In contrast, Gainous (1985) found that teacher attitudes toward giftedness did not affect teacher accuracy in identifying gifted students. Where teacher inaccuracy existed, it was related to teacher knowledge of giftedness. This was corroborated by Schack and Starko (1990), who found that novice teachers tended to base nominations more on grades, classroom performance, and motivation, while expert teachers showed greater preference for IQ scores, vocabulary, and multiple interests. However, Taylor (1986) found that teachers with training in gifted education, who had more knowledge of gifted characteristics, including knowledge of negative manifestations of positive traits, still

Manuscript submitted May, 1993.
Revision accepted December, 1993.