

BY now, most educational leaders have adopted or are aware of a broader concept of giftedness and intellectual talent than has heretofore dominated educational thought. This changing concept has been accompanied by increased interest in the identification of creative talent (Guilford and Merrifield, 1960; Taylor, 1960; Getzels and Jackson, 1962; Torrance, 1962).

Even before the turn of the century, and again in the 1920's and 1930's, interesting and exciting attempts were made to develop tests for identifying creative talent. It was clear from most of this work and the research which accompanied it that such measures tapped different aspects of mental functioning and identified different intellectual talents than did traditional tests of intelligence. Most of these attempts were ignored and there was little interest in the practical use of the instruments developed. Today, however, there is widespread interest in testing for creative talent at all levels of education. Are we ready for this?

There is little question but that research involving recently developed tests of creative thinking identifies a number of kinds of creative talent important in education and society. Many problems remain, however, concerning kinds of creative functioning not tapped by existing tests and the meaning of the measures derived from present procedures. There are problems of further task development, alternative measures of test administration and scoring, and validity. I shall discuss briefly our most recent thinking concerning each of these problems. Our earlier work has been summarized in some detail in *Guiding Creative Talent* (1962).

Testing and Creative Talent¹

We have become increasingly aware of the need for a greater variety of the tasks for assessing the creative thinking abilities by including a greater diversity of stimuli, the involvement of a larger number of the senses, and expanding the limits of the kinds of observations made of task performance. This deficiency becomes quite glaring when work for the past 70 years is viewed in perspective.

In our work at Minnesota we have tended in the direction of complex tasks which have built into them features which make use of what we know about the nature of the creative thinking process, the creative person, conditions which facilitate and inhibit creativity, the qualities of creative products, and the like. We have continued to increase our repertoire of nonverbal tasks. To the Circles Task, we have added the Squares Task, the Parallel Lines Task, various versions of the Incomplete Figures Test, the Pic-

¹ Much of the material included in this article was originally prepared for the Fifth University of Utah Research Conference on the Identification of Creative Scientific Talent, June 6-9, 1962.

E. Paul Torrance is Professor and Director, Bureau of Educational Research, University of Minnesota, Minneapolis.

ture Construction or Shape Test, the Manipulative Design Task, and the Science-Toy Test.

In the verbal tasks, we have used a variety of objects and materials which stimulate the senses of sight, hearing and touch. The Ask-and-Guess Test calls for questions and hypotheses about possible causes and consequences related to a picture. The Product Improvement Test calls for novel ideas for improving such objects as children's toys, objects and instruments used in one's profession, and the like. In most cases, the objects in question have been made available for manipulation. We have now adapted the concept of the well-known Consequences Test by presenting new improbable situations along with drawings and call the resulting tasks, the "Just Suppose Test." The Sound Effects Test makes use of tape-recorded sound effects, a progressive series of warm-ups, and other built-in features based on research findings.

The nature of a test task can be changed by altering the instructions, the nature of the warm-up and motivation, the timing, and other features related to the functioning of the creative thinking abilities. To obtain a maximum of information from a minimum of testing time, an effort has been made to construct batteries or combinations of tasks which embody a variety of these features. A three-task, nonverbal battery provides a good example. The first task, Picture Construction, is designed to stimulate originality and elaboration. The ten-minute time limit is more than most children will use. The high elaborators, however, do not have enough time and some of the highly fluent and highly original individuals keep thinking of ideas to add to or completely change their first idea.

The two succeeding tasks (Incomplete Figures and Circles) increasingly elicit

greater variability in fluency, flexibility, originality and elaboration. There is not enough time to complete all of the possible units and make them elaborate. The Sound Effects Task has been prepared with and without cues or examples as a part of the warm-up and with and without invitations to regress. Versions have also been produced to appeal to imagination as related to all of the senses.

No one has yet made inroads upon the problem of how many tasks, of what length and variety of stimuli are necessary to provide an adequate assessment of creative potential. There is a danger, however, that demands of economy in time and scoring will bring into use tests which represent an oversimplification of sound concepts of creative thinking. There is very real danger in the pressure of some workers to derive a CQ (Creative Quotient). In my opinion, this would be an extremely serious mistake.

Problems of Test Administration

The issues surrounding individual, oral administration versus group, written administration, continue to bother us, especially with tasks requiring multiple verbal responses. Although this problem is more serious in testing children, there are tremendous individual differences even among graduate students in ability to write down one's ideas. At the fourth grade level, scores on group administered tests of creative thinking correlate more highly with scores on tests of intelligence and achievement than scores on individually administered tests. It is doubtful, however, that the group test is the more valid measure of creative talent.

Even at the ninth grade level, there is almost no relationship between scores on tests of intelligence and group measures

of creative thinking within the ranges of the upper 25 percent in intelligence, but there is a relatively high relationship between these two variables within the lower 25 percent in intelligence. These results may occur because the group test of creative thinking does not yield a valid measure of the creative thinking abilities at this level of intelligence. On the Ask-and-Guess Test, the test-retest reliabilities range from .46 to .61 in grades four through six when one test is written and the other oral, whereas the relationships range from .75 to .85 when both administrations are written, even with different stimulus pictures.

Whether or not to give examples as a part of the instructions is an important issue. Present indications are that the giving of examples reduces originality and increases fluency. A similar issue pertains to emphasis upon quantity as opposed to some quality like "clever, unusual and original." In one study, instructions to give "clever, unusual and original" responses actually freed children below the fourth grade to give a larger number of responses. Although the tendency was reversed in the fourth grade and above, the differences are not statistically significant. In another study, when fluency and originality were differentially rewarded, emphasis on originality increased the originality of responses without significantly reducing the fluency.

We are experimenting with different time limits for the same tasks. When time limits have been doubled experimentally, fluency has been greater during the usual time-limit than during the extended time-period. The reverse has been true of originality. We are interested not only in differences in scores on such factors as fluency, flexibility, originality and elaboration, but also in

determining what time limit produces what type of validity.

A similar issue is involved in determining the most desirable length of a battery of tasks to be administered at a single sitting. With batteries of about one hour in length, many individuals reduce both the quantity and quality of their production near the end, while others maintain about the same pace throughout. Knowing this, we must ask, "Does performance on the first tasks yield a better index of potentiality than does performance on the later tasks?" or "Does performance on a lengthy battery administered at one sitting yield a better index of potentiality than does performance on the same battery administered at two or three different sittings?" Perhaps a better question would be, "What kinds of potentiality does each condition enable us to identify?"

Evidences of Validity

In designing test tasks, working out instructions, and developing scoring procedures, we have endeavored to build into them certain characteristics. These characteristics use the best we know about the creative processes as these have been revealed in the history of invention, discovery, and other kinds of creative achievement, as well as the results of experiments, especially those related to deliberate efforts to increase creativity. Our attempts to establish some type of validity have involved one or the other of the following approaches:

1. Identifying high and low groups on some test measure and then determining whether or not they can be differentiated in terms of behavior which can be regarded as "creative."
2. Identifying criterion groups on some behavior regarded as creative and

determining whether or not they can be differentiated from their peers by test scores.

Using the first approach, elementary school pupils who achieved high scores on paper-and-pencil tests of creative thinking also initiated a larger number of ideas in tasks involving creative problem solving than did their less creative peers. When matched for intelligence, sex, race, and teacher, the most creative children, far more frequently than their controls, were found to have reputations for having wild or fantastic ideas, for producing drawings and other products judged as "off the beaten track," and for work characterized by humor, playfulness, and relative lack of rigidity.

Weisberg and Springer (1961), using these same tests with gifted (high IQ) fourth graders, compared the most creative with the least creative. Through psychiatric interviews, the highly creative children were rated higher on: strength of self-image, ease of early recall, humor, availability of Oedipal anxiety, and uneven ego development. On the Rorschach Inkblot Test, they showed a tendency toward unconventional responses, unreal perceptions, and fanciful and imaginative treatment of the blots. Their Rorschach performance described them as being both more sensitive and more independent than their less creative peers.

Using this same approach, a group of high-scoring general business teachers were found to behave dramatically different in their classrooms from a group of their less creative peers. They asked more questions, asked a greater variety of kinds of questions, gave more illustrations of key concepts, and interacted more with their students.

Numerous partial replications of the Getzels and Jackson (1962) studies have

made it clear that tests of creative thinking and tests of intelligence or scholastic aptitude identify different types of individuals. The ways in which they are different describe those high on measures of creative thinking in ways which are usually regarded as "creative."

Using the Minnesota tests, Fleming and Weintraub (1962) obtained a coefficient of correlation of $-.41$ (significant at the .01 level) between six of the verbal tasks and a measure of rigidity, a construct defined in terms of inflexibility, stereotypy, intolerance of ambiguity, and a compulsive need for order.

Using the approach of identifying criterion groups as highly creative or noncreative, children nominated by their teachers on various criteria of creativity, achieved higher scores on the tests of creative thinking than did their peers who were not so nominated. Above the third grade, significant results have also been obtained from peer nominations on creativity criteria. At the college level, indications of validity have come from original-idea projects, scores on subject-matter tests requiring creative problem-solving, self-initiated learning, and from faculty nominations in a technical college where there are numerous opportunities for creative achievements.

Highly productive saleswomen in a large department store (Wallace, 1961) scored significantly higher on a battery of creative thinking tasks than did their less productive peers in the same departments. A sample of mental patients, diagnosed as schizophrenic and judged to be partially recovered, exhibited on the tests extreme inflexibility, frequently blocked, and produced extremely obvious, banal and safe rather than original responses.

In a pilot study, we identified high and low groups on a nonverbal test (Circles)

(Continued on page 72)

Testing

(Continued from page 10)

and studied their language behavior as manifested in imaginative stories. The more creative children wrote longer stories; used a greater number of different words; more frequently used first person pronouns, verbs like "said" and "was," and conjunctions indicating cause or consequence; and the like.

In conclusion, we recognize that the concepts of creative talent which have guided the work described in this article will not satisfy many individuals. Nevertheless, it seems rather clear that high scores based on the measures described identify individuals who behave in ways commonly regarded as creative and that individuals who behave in ways commonly considered as creative achieve higher scores than similar individuals identified as behaving in relatively uncreative ways. Although there is a need for much developmental work, present instruments can be used productively in experimental and research projects.

References

Elyse S. Fleming and S. Weintraub. "Attitudinal Rigidity as a Measure of Creativity in Gifted Children." *Journal of Educational Psychology* 53: 81-85; 1962.

J. W. Getzels and P. W. Jackson, *Creativity and Intelligence*. New York: John Wiley, 1962.

J. P. Guilford and P. R. Merrifield. *The Structure and Intellect Model: Its Uses and Implications*. Los Angeles: University of Southern California, 1960.

C. W. Taylor. "The Creative Individual: A New Portrait of Giftedness." *Educational Leadership* 18: 7-12; 1960.

E. P. Torrance, *Guiding Creative Talent*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

H. R. Wallace. "Creative Thinking: A

Factor in Sales Productivity." *Vocational Guidance Quarterly*, Summer 1961; p. 223-26.

P. S. Weisberg and Kayla J. Springer. "Environmental Factors in Creative Function." *Archives in General Psychiatry* 5: 554-64; 1961.

All Teachers

(Continued from page 29)

mediately measurable. We must place value on the statements of measurement specialists who tell us that the key to improvement of classroom tests is objective analysis of test content and performance in relation to a set of explicit standards for test quality set by the test maker.

All teachers have to make tests and all can be helped to improve—but not without extra effort, extra time, and sufficient money to secure the services of test experts. Critical to the problem is finding the most promising career teachers to involve in the program; teachers who are themselves interested and influential in communication with other teachers. Further, I would argue strongly that an in-service training program on testing must be carried on over periods longer than a year—perhaps two or three. The payoff could only be better teaching.

References

Paul Diederich. "The Problem of Grading Essays." Princeton, New Jersey: Educational Testing Service.

Robert L. Ebel. "The Nature of Modern Tests." Princeton, New Jersey: Educational Testing Service.

"Making the Classroom Test." Educational Testing Service. Evaluation and Advisory Services, Series No. 4. Princeton, New Jersey.

Dorothy A. Wood. *Test Construction*. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961.

Copyright © 1962 by the Association for Supervision and Curriculum Development. All rights reserved.