

Treffinger, D. J., Renzulli, J. S., & Feldhusen, J. F. (1971). Problems in the assessment of creativity. *Journal of Creative Behavior*, 5(2), 104–112.

Problems in the Assessment of Creative Thinking*

**Donald J. Treffinger
Joseph S. Renzulli
John F. Feldhusen**

A large number of studies dealing with various aspects of creative thinking have appeared in educational and psychological publications during the last two decades. Evidence of the extent of this interest appears in Razik's (1965) bibliography and this journal's attempts to update it in 1967 and 1968, as well as in recent volumes of *Psychological Abstracts and Dissertation Abstracts*. Yet there is a great deal of controversy regarding the nature of the creative process and the strategies that hold maximum promise for accelerating creative production. While such controversy is optimistically viewed as a healthy symptom in any relatively new line of scientific inquiry, our failure to master certain basic problems after nearly twenty years of intensified study has led to a decrease in interest among educational practitioners who at one time were eager to rally round the flag of creativity and to "do something" about this newly discovered (or rediscovered) human ability. Unless researchers can begin to find answers to many unsolved problems, the concept of creativity may be, at best, a catchall. At worst, there exists a very real danger that it could eventually be tossed upon the junk heap of discarded educational fads.

The purpose of this paper is to provide an overview of the major problems and issues that relate to the scientific study of creativity. By isolating the important dimensions of the problem, we hope that some direction may be provided for future research efforts.

Assumptions of This Paper

There are two basic underlying assumptions upon which this paper is based. The first is that certain unique psychological processes, referred to as "creativity," do in fact exist in man's repertoire of behaviors, although in our investigation of those behaviors, we may have merely scratched the surface. The second assumption is that creative process is complex, or multidimensional, in nature.

* This paper is based on the authors' presentations in the symposium Assessing Creativity: Progress in Both Directions; at the Annual Meeting of the American Educational Research Association, Minneapolis, March, 1970, Richard E. Ripple, Chairman.

Problems Treated

In this paper, we will consider two general and interrelated problems and several specific issues within each. The first set of problems involves the *theoretical description* of creative thinking; the second will be referred to as the *criterion problem*.

Theoretical Description of Creative Thinking

The first problem, then, is that there is no single, widely-accepted theory of creativity which can serve to unify and direct our efforts at specifying an adequate assessment procedure. The work of Mednick (1962) and his associates illustrates, perhaps as well as any, the formulation of a theory of creativity from which a particular method of assessment emerges. Yet, for a number of reasons, many researchers have not been attracted to this theory (cf. Cropley, 1966; Jackson & Messick, 1965; Taft & Rossiter, 1963), and it can hardly be described as widely accepted. Other theories, such as those of Rogers (1962) or Kubie (1958), have seldom resulted in the formulation of psychometrically adequate assessment procedures. Guilford's widely-known "structure of intellect" model (1967) does not constitute a theory of creativity *per se*, despite the fact that it has been heuristically or conceptually useful in describing some cognitive abilities which are related to creativity. It may be useful to describe it as a theory of human intelligence which subsumes some important cognitive aspects of creativity. Even though Guilford (1967) has argued in recent discussions that creative thinking is not merely a matter of divergent production, a comprehensive theory of creativity would necessarily consider in detail the nature and interrelationships of non-cognitive components of creative behavior, as well as the cognitive aspects.

Torrance's (1966) tests purport to be broadly eclectic, drawing from the "best" of theory available at the time of their development, but for that very reason—that they lack a unified, comprehensive, theoretical base—difficulties are inevitable. Of course, the variables assessed by the Torrance Tests (fluency, flexibility, originality and elaboration) are all classified in Guilford's "structure of intellect."

The Problems of Measurement

Given the existing array of ideas about creativity, and the absence of "theoretical unity," it is not in the least surprising that there exists a number of tests, all purporting to be measures of "creativity," but differing in a number of ways. Each instrument mirrors the particular set of beliefs and preconceptions of its developer concerning the nature of creativity. Sadly, the theoretical rationale for such tests is often not even sufficient to allow systematic tests of differential predictions.

An outgrowth of this problem, although a major concern in its own right, is that *we do not understand very completely the implications of differences in assessment procedures*. Variations in working time, test atmosphere, and directions given to the examinee, for example, seem to yield different kinds of results and different patterns of intercorrelations between creativity scores and other cognitive or achievement variables. It is quite clear that such changes occur (Van Mondfrans, Feldhusen,

Treffinger, & Ferris, in press; Wallach & Kogan, 1965). What is *not* clear is the reason for those changes, or under what conditions certain results might be predicted.

Van Mondfrans et al. (in press) argued that the matter is much more complex than merely removing the time limits and appearances of a test-like situation. Removing time limits, for example, had no significant effect on pupil performance on verbal tasks. The highest scores on these tasks were obtained under standard “test-like” conditions. On figural tasks, however, removing time limits did influence pupil performance; highest scores were obtained by pupils under “take home” conditions.

Continuing experimental work is needed to understand the problems of test procedures and their implications more completely. Such research would also be more profitable if predictions could be derived from a specific theoretical conception of creativity. In the meantime, a clear implication seems to be that researchers who use “creativity tests” should be extremely careful to report in detail the procedures for test administration, directions, and timing.

Creativity's Relation to Other Abilities

Another very controversial issue, which is related to theoretical problems and has probably prevented educators from achieving some closure in programming for the classroom, is a problem which we will refer to as *dimensionality*. (In measurement terms, the issue is more properly referred to as convergent and discriminant validation; see Campbell and Fiske, 1959). Simply stated, the dimensionality issue involves the degree to which measures of creativity or divergent thinking are empirically distinguishable from other more traditional measures of cognitive processes such as intelligence and academic achievement. The development of defensible measures of creativity would seem to depend on constructing a series of tasks which share substantial variance with each other, but are at the same time generally independent of other traditional cognitive measures. The concern for this problem is reflected in the disproportionate amount of research that has been devoted to the creativity-intelligence distinction and our inability to arrive upon a generally acceptable operational definition. (Taylor, 1959, for example, has listed over one hundred definitions which have added to the semantic fog that envelops the study of creativity.) A great deal of the concern for the dimensionality issue, and the lack of resolution of this issue, stems from the problem of measurement and the adequacy of currently available tests of creativity and the divergent-thinking processes.

A number of research studies (Ripple & May, 1962; Thorndike, 1962; Wallach & Kogan, 1965) have cautioned against the uncritical acceptance of the Getzels and Jackson (1962) hypothesis which suggested that creativity and intelligence were unrelated. In a historical perspective upon the measurement of cognitive processes, Ward (1963) called attention to aspects of Binet's and Wechsler's classic definitions of intelligence, parts of which sound surprisingly similar to many present-day definitions of creativity. Others (Guilford, 1967; Wallach, in press) have made a similar case for the relationship between creativity and the classic definitions of problem-solving.

As a result of the lack of a unified, widely-accepted theory of creativity, then, educators have been confronted with several difficulties: establishing a useful operational definition, understanding the implications of differences among tests and test administration procedures, and understanding the relationships of creativity to other human abilities.

The Criterion Problem

The second general problem has been described as the *criterion problem*. What criteria exist against which the validity of creativity tests may be assessed? Although this problem has not generated as much concern as the creativity-intelligence controversy, its interrelatedness to all other aspects of the study of creativity demands that it be given high priority among areas in which research is needed.

Many researchers have tended, on the one hand, to view creativity entirely as a cognitive process, or, on the other hand, entirely as a complex set of personality traits. The former have tended to ignore the possibility that there may be an affective component to creativity, and the latter have tended to overlook the importance of underlying cognitive abilities in creative problem-solving. It is most likely, however, that a valid assessment procedure would, of necessity, consider both components. In the meantime, we must be very cautious about our willingness to make inferences about "creativity" from measures which are distinctly cognitive, particularly the divergent-thinking-type tests. This does not imply rejection of the usefulness of tests of divergent thinking. It may be true that some of the critics have been too severe (e.g. Covington, 1968; Wallach, 1968). While divergent-thinking measures certainly do not tell the entire story about creativity, it is quite likely that these measures do assess intellectual abilities which play an important role in creativity. If creativity is viewed as a complex kind of human problem-solving (in which case perhaps the term "creative problem-solving" would be preferable), divergent thinking may be a necessary, although not a sufficient, component.

Teacher and Peer Judgments

There have been many difficulties in identifying acceptable external criteria for the validation of creativity tests. Foremost among them is the difficulty of any attempt to use teacher and peer judgments as a means of identifying creative youngsters. A number of studies which sought to use this approach (Holland, 1959; Reid, King, & Wickwire, 1959; Rivlin, 1959; Torrance, 1966; Wallen & Stevenson, 1960; Yamamoto, 1964) have shown that when teachers and peers are asked to nominate very creative pupils or those with good imagination or many new ideas and ways of doing things, they usually produce a list of classmates who are the highest achievers or have the highest IQs. Further, there is considerable variability among teachers in the ability to rate pupils against a test criterion, even when specific definitions are provided. Research is needed on the effectiveness of procedures for training teachers or peers to be more effective raters, less influenced by other criteria.

Creativity Profiles

Related attempts to establish external criteria for creativity have been the well-known series of studies that analyzed the characteristics of adults who have made significant contributions to their respective professions (e.g., Barron, 1969; MacKinnon, 1962). While these studies have provided us with excellent profiles and descriptions of the highly creative person, we must be careful not to confuse concurrent validity with predictive validity. MacKinnon (1962) cautioned that it is one thing to discover distinguishing characteristics, but quite another matter to conclude that traits observed several years after school or college truly characterized an individual when he was a student. Nor can we conclude that these same traits in youngsters today will identify individuals with the kind of creative potential that will be valued in tomorrow's world.

Products as Indices of Creativity

Another approach to the criterion problem would be to use products as indices of creative achievement. Thus great discoveries, inventions, works of art, or writings could be used as criteria. In the research by MacKinnon (1962) and Barron (1969), such indices were undoubtedly often used as the basis for judging an individual's significant contribution in a field. Miles (1968) has also attempted to develop tasks for concurrent assessment of an individual's ability to produce a creative object. While seemingly a hopeful approach to the development of criterion measures for validation of creativity tests, this approach through the use of products is beset by reliability problems.

Problem-Solving Tasks

There is also some reason to believe that some of the problems of assessing creative problem-solving relate to the heterogeneity and underdevelopment of the tasks that have been employed. As Davis (1966) and others have pointed out, the literature on problem-solving is very confusing. "Creative problem-solving" tasks have been used in one study and then never used again. Some people have attempted, as Davis did, to categorize or classify problem-solving, but this classification has tended to be rational rather than empirical. Some logical groupings or judgments about tasks may not hold up very well under closer examination; tasks which "on the face" seem to be attractive measures of creative problem-solving may reflect quite different appearances when studied empirically.

There is a great deal to be learned about the assessment of creative problem-solving. It is quite clear that simple measures of fluency, flexibility, and originality are not sufficient. Perhaps substantial effort must be given to finding new, more complex measures. Perhaps as a beginning we must at least look more carefully at the *interactions* among divergent-thinking scores (fluency-flexibility interactions, for example) and between divergent-thinking scores and other abilities; very little use of such combined subscores seems to have been made in the literature.

There are also a number of problems of a very practical nature to solve. How does the researcher know that what *he* considers creative tasks are creative and challenging for the examinee? It may be that the tasks he considers most unusual are boring, unexciting, even trivial, for the most imaginative of our examinees. Perhaps each task that purports to be an assessment of creative problem-solving should be accompanied by a simple rating scale: "Have you ever worked on this problem before? Did you solve it? Were you given the solution? What did you think about the problems you have solved here? Were they interesting? Challenging? What did you think of your solutions?" Although many psychologists avoid using the term "introspection," it may be that quite a bit could be learned about measures of creativity by asking subjects to talk about their experiences. Perhaps the adequacy or creative strength of a response, or the extent to which a task captures the subject's attention and stimulates him to think creatively, are important matters, but only capable of being assessed by the subject himself.

Measures of Originality

Another dimension of the criterion problem concerns the appropriateness or inappropriateness of our current means for assessing originality. While a few have dissented, almost everyone who has grappled with creativity research appears to be satisfied with the statistical infrequency criterion for measures of originality. At least one researcher (Starkweather, 1964, 1968) has attempted to devise an alternate procedure, involving comparison of a child's response to all of his *own* responses, rather than to the responses of other children. Perhaps our easy acceptance of the statistical infrequency criterion has prevented us from identifying new methods which are useful for measuring this dimension of creative thinking. Ideally one would like a qualitative index with face validity.

An issue of critical importance in solving the problem of assessing creative thinking is concerned with the validity of our measures. Too often, in order to develop tests which are manageable from the psychometric point of view, we have relied on tasks which may have little or no logical relationship to creative behavior as it occurs in the "real world." While there exists a substantial difficulty (identifying adequate criteria against which the test tasks can be validated), the problem warrants our attention. The "creativity" assessed by our tests, after all, should be expected to bear a resemblance to creativity as it is actually manifested among people. Finally, we should at least acknowledge the existence of a number of other important issues in research on creativity and its assessment: assessing the relevance of responses, distinguishing between sensible and bizarre responses, and establishing differential age and sex criteria. Most would agree that these are essentially unresolved problems, and thus appear to be topics that are worthy of the researcher's attention. Occasionally, the study of creativity has been described as a classic case of the blind leading the blind, but researchers in this area may prefer to look upon the situation as somewhat of a challenge, and to keep in mind that in the land of the blind, a one-eyed man can be king!

References

- Barron, F. (1969). *Creative person and creative process*. New York: Holt, Rinehart & Winston.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105.
<https://psycnet.apa.org/doi/10.1037/h0046016>
- Covington, M. V. (1968). *New directions in the appraisal of creative thinking*. Berkeley, CA: University of California, unpublished mimeograph.
- Cropley, A. J. (1966). Creativity and intelligence. *British Journal of Educational Psychology*, 36(3), 259–266. <https://doi.org/10.1111/j.2044-8279.1966.tb01878.x>
- Davis, G. A. (1966). Current status of research and theory in human problem-solving. *Psychological Bulletin*, 66(1), 36–54. <https://doi.org/10.1037/h0023460>
- Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students*. New York: Wiley and Sons.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Holland, J. L. (1959). Some limitations of teacher ratings as predictors of creativity. *Journal of Educational Psychology*, 50(5), 219–223.
<https://psycnet.apa.org/doi/10.1037/h0040598>
- Jackson, P., & Messick, S. (1965). The person, the product and the response: Conceptual problems in the assessment of creativity. *Journal of Personality*, 33(3), 309–329. <https://doi.org/10.1111/j.1467-6494.1965.tb01389.x>
- Kubie, L. S. (1958). *Neurotic distortion of the creative process*. Lawrence, KS: University of Kansas Press.
- MacKinnon, D. W. (1962). The nature and nurture of creative talent. *American Psychologist*, 17(7), 484–495. <https://psycnet.apa.org/doi/10.1037/h0046541>
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220–232. <https://psycnet.apa.org/doi/10.1037/h0048850>
- Miles, D. T. (1968). *Development of a test for an experimental research program in creative problem solving. Final report* (Bureau No. BR-7-E-037). Washington, DC: U.S. Office of Education, Bureau of Research.
<https://files.eric.ed.gov/fulltext/ED024208.pdf>
- Razik, T. (1965). *Bibliography of creativity studies and related areas*. Buffalo, NY: Creative Education Foundation.
- Reid, J. B., King, F. J., & Wickwire, P. (1959). Cognitive and other personality characteristics of creative children. *Psychological Reports*, 5, 729–737.
<https://psycnet.apa.org/doi/10.2466/PRO.5.7.729-737>
- Ripples, R. E., & May, R. B. (1962). Caution in comparing creativity and IQ. *Psychological Reports*, 10(1), 229–230.
<https://doi.org/10.2466/pr0.1962.10.1.229>
- Rivlin, L. G. (1959). Creativity and the self-attitudes and sociability of high school students. *Journal of Educational Psychology*, 50(4), 147–152.
<https://psycnet.apa.org/doi/10.1037/h0045328>
- Rogers, C. R. (1962). Toward a theory of creativity. In S. J. Parnes & H. F. Harding, (Eds.), *A source book for creative thinking* (pp. 64–72). New York: Scribner's.

- Starkweather, E. K. (1964). Problems in the measurement of creativity in preschool children. *Journal of Educational Measurement*, 1(2), 109–133.
<https://www.jstor.org/stable/1433677>
- Starkweather, E. K. (1968). Studies of the creative potential of young children. In F. E. Williams (Ed.), *Creativity at home and in school* (pp. 75–122). Saint Paul, MN: Macalester College.
- Tait, R., & Rossiter, J. (1966). The remote associates test: Divergent or convergent thinking? *Psychological Reports*, 19(3), 1313–1314.
<https://doi.org/10.2466/pr0.1966.19.3f.1313>
- Taylor, I. A. (1959). The nature of the creative process. In P. Smith (Ed.), *Creativity*. New York: Hastings House.
- Thorndike, R. L. (1962). *Some methodological issues in the study of creativity*. In Proceedings of the 1962 invitational conference on testing problems. Princeton, NJ: Educational Testing Service.
- Torrance, E. P. (1966). *Torrance tests of creative thinking: Norms-technical manual*. Princeton, NJ: Personnel Press.
- Van Mondfrans, A. P., Feldhusen, J. F., Treffinger, D. & Ferris, D. (in press). The effects of instructions and working time on divergent thinking scores. *Psychology in the Schools*.
- Wallach, M. A. (1968). Review of the Torrance Tests of Creative Thinking. *American Educational Research Journal*, 5(2), 272–281.
<https://doi.org/10.3102/00028312005002272>
- Wallach, M. A. (in press). Creativity and the expression of possibilities. In S. R. Graubard, & J. Kagan (Eds.), *Creativity and Learning*. Boston: Houghton Mifflin.
- Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity-intelligence distinction*. New York, NY: Holt, Rinehart, and Winston.
- Wallen, N. E., & Stevenson, G. M. (1960). Stability and correlates of judged creativity in fifth grade writings. *Journal of Educational Psychology*, 51(5), 273–276.
<https://psycnet.apa.org/doi/10.1037/h0045933>
- Ward, V. S. (1963, May 2–4). *Developing productive thinking: Educational implications*. Paper presented at the Second Conference on Productive Thinking, National Education Association, Washington, DC.
- Yamamoto, K. (1964). Evaluation of some creativity measures in a high school with peer nominations as criteria. *Journal of Psychology*, 58(2), 285–293.
<https://psycnet.apa.org/doi/10.1080/00223980.1964.9916746>

Donald J. Treffinger, Associate Professor, Educational Psychology. Address: Purdue University, School of Humanities, Social Science and Education, Lafayette, Indiana.

Joseph S. Renzulli, Associate Professor, Educational Psychology. Address: University of Connecticut, Storrs, Connecticut.

John F. Feldhusen, Professor and Chairman, Educational Psychology Section. Purdue University, Lafayette, Indiana.