DOCUMENT RESUME

ED 374 140 TM 022 016

AUTHOR	Venable, B.
TITLE	A Philosophical Analysis of Creativity Measurement.
PUB DATE	27 Apr 94
NOTE	24p.
PUB TYPE	Information Analyses (070) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Attitude Measures; Convergent Thinking; Creative Thinking; *Creativity; Creativity Tests; Definitions; Divergent Thinking; *Educational Philosophy; Elementary Secondary Education; Evaluation Methods; Intelligence; Interest Inventories; *Measurement Techniques; Test Construction; *Test Use
IDENTIFIERS	*Self Report Measures

ABSTRACT

Whether the aim has been to understand the nature of creativity or to determine its development, researchers have formulated a number of measures of creativity. A glut of results and conclusions actually hinders an educational understanding of the term creativity. The most widely known tests of creativity incorporate divergent-thinking, as opposed to convergent- thinking, skills. Association tests provide another way to measure creativity, as do attitude and interest scales. Creativity has been regarded as an aspect of personality and has been measured through biographical inventories. Yet another form of measurement is through nomination by teachers, peers, or supervisors. The assessment of products created is another way creativity has been measured. Studies of eminent persons and self-reports are other ways to examine creativity. The variety of approaches reaffirms that no definite conclusion exists about how creativity should be measured, and the relations between creativity and intellect remain unclear. The attempt to label children as creative or as having creative ability is premature, in that there is no general definition of creativity. Research must clarify ambiguities of definition and the role of creativity in education. (Contains 13 references.) (SLD)

A Philosophical Analysis of Creativity Measurement

by B. Venable
April 27, 1994

A Philosophical Analysis of

Creativity Measurement

There is a goal in education to derive meaning from its complex nature. Measurement in the form of testing is one example of this attempt that has evolved in the area of creativity. This paper attempts to better understand this phenomenon and its relationship with our educational system.

Introduction

Assuming that creativity is an aspect of the intellect that to some degree is involved in human development, it has been the educational researcher's aim to gain insight into its complexity. Whether the aim has been to understand the very nature of creativity or ascertain its development, or to identify those who have it and to what degree, researchers have formulated and engaged numerous measurement devices to accomplish this task.

The Problem

Although few would argue that creativity is worthwhile, little consensus exists as to what aspects of this phenomenon need measuring. Further, the enormous offerings of current assessment tools and the accumulated findings have resulted in a complexity of outcomes that hinder creativity's role in education (Crockenberg, 1972).

I am reminded of a metaphor in which several blind-folded people are situated around an elephant, each touching some aspect of the animal. The ensuing individual definitions of 'elephant' from divergent vantage points only shed light on a small part of the large whole. In the case of creativity testing, researchers have developed such a plethora of methods that there exists a glut of complex results and conclusions, many inconclusive, rendering this animal called creativity educationally impotent.

Major Thesis

The following information lends credence to the problem. Through offering a history and overview of measurement devices, and providing the reader with discussion and criticism associated with the findings, implications about creativity and its measurement will be made.

Early Research

Some of the first research including creativity is associated with the techniques employed by Binet and Henri in 1896 using inkblot interpretations, and measurement of subjects who accomplished a theme or drawing, and completed a sentence using given words (cited in Wakefield, 1991).

Later in 1915, Whipple, in a manual of tests for complex processes, included instruments for measuring

"imagination and invention" (cited in Torrance & Hall, 1980). Although reported to have caused some controversy, this, like most earlier work was generally ignored by the research community of the time.

Some attention was given to the work of Johnson O'Connor who in 1940 and 1948, created a test which went in the direction of the "...further reaches of creative potential." (cited in Torrance & Hall, 1980). Several volumes of research were published by the Human Engineering Laboratories who used this test. Terms such as "ideaphoria", "creative imagination", and "visual imagination" were visages of this test.

Divergent Thinking Tests

Probably the most widely known and used tests of creativity incorporate divergent thinking skills as opposed to convergent thinking skills. The difference between them was once considered the distinction between creativity and intelligence (Wakefield, 1991). J.P. Guilford's structure of the intellect (1956) was the basis for many tests based on this concept (cited in Hocevar, 1979). They differ from traditional intelligence tests in that they reward a multitude of answers rather than a single correct one.

A number of other researchers have modified the ideas of Guilford and incorporated them into creative assessment tools of their own. Some of them include

Getzels & Jackson (1962), Gough (1975, 1976), Industrial Relation's Center (1959), Mednick, S. & Mednick, M. (1967), and Wallach & Kogan (1965).

The most frequently used test of creativity is the Torrance Test of Creative Thinking (1966, 1990) which employs divergent thinking as an assessment tool (cited in Baer, 1994; Crockenberg, 1972; Treffinger & Poggio, 1972; Davis, 1975; Wakefield, 1991; Hocevar, 1979). In this and other divergent models creativity is usually defined in terms of fluency (lots of ideas), flexibility (many different ideas), originality (unique ideas), and elaboration (detailed ideas). To illustrate how popular this test has become, a cumulative bibliography on the Torrance Test of Creative Thinking listed over 1,000 publications that involved this testing instrument (cited in Torrance & Hall, 1980).

The 1962 Wallach and Kogan Creativity Test (cited in Crockenberg, 1972) was similar to that of Torrance. Both tests rewarded divergent answers and had a combination of verbal and visual content. However, the Wallach and Kogan Creativity Test differed from Torrance in its relaxed testing atmosphere, which permitted fewer time constraints, favoring a game-like rather than strict examination-type setting.

Association Tests

The Remote Association Test (1967) from the

researchers Mednick and Mednick offers still another method to assess creativity (cited in Crockenberg, 1972; Hocevar, 1979; Davis & Belcher, 1971). In this test, individuals respond to 30 convergent-type items where three given words are presented and a fourth word must be found that relates to all three. For example, *surprise*, *birthday*, and *line* are given. *Party* might be an associated answer. This defines the creative individual as one with a "...steep associative hierarchy..." (Davis & Belcher, 1971).

Attitude and Interest Scales

Another testing device derives its information from attitude and interest inventories. The use of this test follows the assumption that creative persons will express interest in creative activities (cited in Hocevar, 1979). In an inventory by Taft and Gilchrist (1970), a creative individual would select the following items:

Invent a new gadget.

Think up new plots.

Write words in new combinations to convey emotion.

Still another inventory by Baird and Holland (1968) would ask subjects having a high degree of originality to agree to items such as the following:

I often daydream about unsolved problems.

I often act without thinking.

I would rather be a research engineer than an industrial engineer.

I have to learn things in my own way rather than accepting ideas or relationships suggested in textbooks, etc.

Other attitude and interest inventories (cited in Hocevar, 1979) are the Runner Studies of Attitudinal Patterns (1954), the Covington Attitudinal Inventory for Problem Solving (1966), the Creative Behavior Disposition Scale (1974), the Creative Attitude Survey (1970), and the Opinion, Attitude, and Interest Survey (1965).

Personality Scales

Creativity is often seen by investigators as an aspect of personality. Established personality inventories are scaled to identify creative abilities (cited in Hocevar, 1979). One such tool is the Adjective Check List which uses terms like "clever", "complicated", "cynical", "imaginative", "original", and "reflective" as representative adjectives of creative individuals.

Other personality inventories which have been modified to test for creativity are the Omnibus Personality Inventory (1968), Gough's California Psychological Inventory (1957), Cattell and Eber's Sixteen Personality Factor Questionnaire (1968), and

Barron's Independence of Judgment Scale (1953). Finally, Torrance and Khatena (1970) designed a personality inventory to identify creativity in adolescents. In this study, entitled "What Kind of Person Are You?", the subject would choose characteristics from forced-choice offerings such as curious rather than self-confident, or altruistic as opposed to courteous.

Biographical Inventories

Biographical Inventories make up another distinct category of creativity tests (cited in Hocevar, 1979; Davis, 1975). Assuming that creative behaviors are somehow related to past influences, researchers have attempted to predict creative talent by studying an individual's past.

The Alpha Biographical Inventory is one such tool created by the Institute for Behavioral Research in Creativity (1968). It is the result of extensive research conducted on NASA scientists and engineers and is particularly applicable to creativity in the sciences. Schaefer's Biographical Inventory (1970) has much broader applicability. Questions from this inventory are grouped in sections of physical characteristics, family history, educational history, leisure time activities, and a miscellaneous category.

Nominations

The actual educational setting provides an excellent environment for research. Some researchers have discovered that nominations by either teachers or peers is yet another form of creativity measurement (cited in Hovecar, 1979).

In the area of teacher nominations, Yamamoto's work (1963) at the elementary level is an example. Using selection methods of fluency, flexibility, inventiveness, and originality, teachers were asked to nominate students who they thought were the most creative. Nelson (1963), on the other hand used a checklist of personality traits for teachers to use in nominating creative students.

At the secondary level, Rivlin (1959) had 25 teachers attend a training conference where 14 standards of creativity were discussed. On the basis of these standards teachers were asked to nominate five intelligent and creative students and five equally non-creative students as well.

At the higher education level, Drevdahl (1956) asked faculty members from an arts and science department to nominate creative students using a seven-point scale of creativity. In 1976, Getzels and Csikzentmihalyi (cited in Wakefield, 1991; Hocevar, 1979) studied the problem at the Art Institute of

Chicago. Faculty members were asked to rate both the student's originality and artistic potential. Wakefield's review of this study found it to be, "To date, the best study of problem finding..." This study distinguished between the act of problem solving and problem finding by asking students to choose from an array of objects from which they would do a drawing. This allowed them to manipulate type, number, and difficulty of object(s) selected.

Peer selection represents another method of nomination for creativity. In 1962, Torrance (cited in Hocevar, 1979) attempted to clear up the inherent ambiguities of such research with children by using criteria categories. These criteria, he believed, would aid in the evaluations of those who were to nominate peers with creative abilities. Other researchers who have developed peer nominating devices are Yamamoto (1964), Foster (1971), and Reid, King and Wickware (1959).

Although most nomination tools are used with children, one study by Taylor, Smith and Ghiselin (1963) used this technique to identify creative engineers (cited in Hocevar, 1979).

Supervisor Nominations

Identifying employees with creative abilities has been the topic of several researchers (cited in Hocevar,

1979). Using characteristics of creativity ascertained from research supervisors, Buel (1960) developed a checklist which can be used to help identify creative research personnel. Taylor (1963) and others working at the Institute of Personality Assessment and Research have also developed a rating scale for identifying creativity in scientists working at a government research lab. Included in this study were traits such as productivity, drive, mathematical ability, integrity, desire for facts, flexibility, cooperation, persistence, and creation.

Product Assessment

The assumption that creative people produce creative products is the foundation of a test developed by Foster in 1971 (cited in Hocevar, 1979). His rating scale judged activities normally associated with secondary schools. Using guidelines based on fluency, flexibility, and originality, expert and non-expert judges evaluated such products as story writing, model building, playing charades with a given set of props, and mathematical equalities with given sets of numbers and symbols.

Studies of Eminence

The Institute of Personality Assessment and Research (cited in Hocevar, 1979) conducted a study to ask architects, writers, and mathematicians to select

the most creative people in their respective fields. Those selected were asked to visit Berkley and participate in a study by IPAR. Similar studies have attempted to identify men (and women?) who also fit this category of eminence. Cattell (1903), Ellis (1904), Roe (1951), and Barron (1969) have all participated in research of this type using a wide range of categories and fields.

Self-Reported Creativity

Creative activities are measured by individuals in our society in any number of fields. In 1964, Holland and Nichols described lists of creative activities they considered rare and meaningful (cited in Hocevar, 1979). This list was based on the study of talent used for the National Merit Scholarship Cooperation which lists achievements in art, science, literature, music, etc. Placing first in a competition, publishing a poem or story, having a dramatic role which is produced at a college or university, or inventing a patentable device are some of the checklist items of this study. Other lists similar in content have been developed by Erickson (1966), Hocevar (1977), Torrance (1969), Skager, Schultz & Klein (1965), and Wallach and Wing (1969).

Miscellaneous Studies

By no means has the list of creativity tests and assessment tools been exhausted here. Several defy a

general categorization because they are quite specific in what creative qualities they are attempting to measure (cited in Hocevar, 1979). They include the Starkweather Creativity Test (1971), the Ideal Pupil Checklist (1975), the Barron-Welsh Art Scale (1952), the Welsh Figure Preference Test (1959), the Pennsylvania Assessment of Creativity Tendency (1971), the Group Inventory for Finding Creative Talent (1976), the Ingenuity Test (1968), and the Onomatopoeia and Images (1969).

Discussion and Criticism

The sheer number of creativity tests and the complex nature of their results prohibits a thorough examination of their value here. It is hoped that a general critique and discussion of them will provide the reader with an appreciation of the philosophical problem inherent in creativity measurement.

Definition of Terms

Reliability, *validity*, and *usability* are three general categories which may be employed in the assessment of psychological tests (Treffinger & Poggio, 1972).

Reliability is "The extent to which a measure is free from random error." (Harre & Lamb, 1983). Treffinger and Poggio (1972) have stated that reliability has three aspects. They are: *stability*

which refer to scores remaining similar over a period of time, *equivalence* or *comparability* which assesses reliability in the administration of alternate tests, and *internal consistency* which assumes that performance on one part of the test will not differ greatly from performance on another.

Validity, as defined by Harre and Lamb (1983) is "The extent to which scores on tests or other measures are justified or supported by the evidence." Validity is measured in three aspects: *content validity* which is "the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be sampled" (Anastasi, 1968), *criterion-related validity* which involves predicting an individual's behavior in a specific situation, and *construct validity* which attempts to validate a test's ability to measure the theoretical construct or trait (Treffinger & Poggio,1972).

Usability refers to the practical consideration and use of the testing mechanism. Primarily, the aspects involved in usability are test administration, test scoring, and norms (Treffinger & Poggio, 1972).

It is important that the aforementioned categories are understood as these aspects of measurement pose particular problems to researchers of creativity.

The Multi-variant/Uni-variant Debate

There are several problems that are associated with creativity tests as they relate to content validity. Reading indicates that foremost among them is the idea that creativity is a single and measurable trait. Although most tests tend to assume this definition, there seems to be growing research that creativity is a complex set of many traits. Tests that are limited to divergent thinking skills, for example "...cannot be content valid as an assessment of creativity, since it is known to sample only a small portion of the abilities which contribute to creative talent." (Treffinger & Poggio, 1972). Although Guilford (1971) disagrees that divergent tests have validity, he considers the idea of creative talent as a single, distinct variable to be erroneous. Baer (1994) reports that studies show, "...cognitive abilities underlying creative performance differ from task to task." Further, creativity measurement rarely explores the diverse nature of its expression. "There is little differentiation among creativity in the arts, sciences, literature, etc. Yet, intuitively it is plausible that a person who is creative in one area has neither the time, ability, nor the motivation to be creative in other areas." (Hocevar, 1979).

Another difficulty relating to a multi-variant

exploration of creativity is the selection of criteria. Treffinger and Poggio (1972) ask, "What are the external criteria by which measures of creativity may be validated?" Here, establishing criterion-related validity becomes difficult because of the lack of concensus as to what 'novelty', 'uniqueness', or 'creative products' are for example. This casts doubt on tests that involve checklists and product judgment as criterion for evaluation.

General Theory Criticism

Another content validity problem is the "...absence of a simple generally accepted theory of creativity which would serve to unify or direct efforts at specifying assessment procedures." (Treffinger & Poggio, 1972). An all-encompassing theory would make the testing and training easier, as well as more cost effective, but research indicates no general theory will be embraced in the near future (Baer, 1994).

Other obstacles exist within the aspect of construct validity. Because there is no generally accepted creative theory, there is difficulty in formulating a testable hypotheses or a rationale for a hypotheses (Treffinger & Poggio, 1972). The development of the hypotheses is "essential" to construct validity according to The American Psychological Association's *Standards for Educational and Psychological Tests and*

Manuals (cited in Treffinger & Poggio, 1972).

Relationship with IQ Test Scores

It is unclear what relationship IQ scores have to creativity. Researcher's attempts to correlate the two have met with limited success (Crockenberg,1972). The controversy over this relationship is believed by Treffinger and Poggio (1972) to relate to "...problems in the definition and theoretical interpretation of both creativity and intelligence." Crockenberg (1972) has noted that there is "...substantial evidence to indicate that there might be a threshold of intelligence above which there is little relationship between IQ and creativity." In general, a fairly high IQ may be required in order for creativity to occur, but beyond that requirement there seems to be no relationship between the two.

Reliability

Treffinger and Poggio (1972) suggest that stability in creativity tests has not provided "clear evidence for automatic acceptance" of reliability. They point to several theoretical possibilities for this:

1. The determination as to whether creativity is, in fact a stable human characteristic.

2. The identification of the motivational influences of the subject tested.

3. The identification of an appropriate interval

between test and retest requirements of some assessment devices.

4. The determination that the measurement universe may be a partial or incomplete sampling.

With regard to internal consistency as a reliability measure, it is unclear whether non-traditional testing methods, especially ones that are open-ended in their answer formats, can be evaluated under this criteria (Treffinger & Poggio, 1972).

Usability Problems

Test scoring which is subjective in nature tends to develop problems with the criteria used in the assessment process and in the judgment of originality and imagination (Treffinger & Poggio, 1972).

Establishment of norms too, presents certain problems. Treffinger and Poggio (1972) report that there are those who feel that the very nature of creativity prevents a development of normal distribution. That particular view point is seen by the authors as removing creativity out of the domain of most normal individuals and placing it in the realm of the rare and exceptional.

General Criticism

Reading indicates that current tests of creativity shed light not upon those who are creative, but upon the complex nature of creativity itself. Hocevar (1979)

reports that, "Any study of creativity is severely limited by its assessment method." He goes on to state, "Tests of cognition, attitudes, interests, personality, biography, etc. are appropriate when the researcher's goals are to explain something about creativity, but they are not acceptable in the selection process. High scorers on these tests are not necessarily creative people." He concludes by reporting that, "...peer nominations, supervisor ratings, teacher nominations, and judgment of products are inadequate indicators of creativity due to the rater's inability to discriminate creativity from other traits. Divergent thinking, biographical characteristics and personality characteristics are best stated as correlates of real life creative behavior, and they should not be taken as direct measures of creativity." Much research substantiates this opinion to some degree.

Hocevar continues to report that studies involving eminent people were felt to be impractical and lists of their accomplishments too rare and exclusive to be much help to the general population. Conversely, Hocevar's highest evaluation of creativity assessment was in the area of self-reported creativity. He found that it was superior in the measure of psychological traits and a good predictor of future behavior.

Summary

Creativity is a vastly complex phenomenon. In the drive to gain an understanding, researchers have developed a variety of testing devices to pin down this concept. Creativity, however, has remained elusive. In fact, this animal has grown even more complex, due to persistent, yet inconclusive discussion and the results of copious assessment devices circulating around it.

Implications and Conclusion

The term *measure* has several facets worth noting. It is defined in terms of "means", "size or quantity", "dimension", "unit", and "proportion" (Simpson & Weiner, 1989). Common to all these descriptors is the notion of a concrete result, and one that can withstand assaults of ambiguity and vagueness. But it seems to be the very nature of creativity that it has remained ambiguous and vague in its processes and definition. Research-driven study with its myriad of results has frustrated the observer. While no definitive conclusion exists about how creativity should be measured or how those results will be used in schools, there is movement to make some gifted and talented decisions based in part on some of these creativity test results (Crockenberg, 1972). Creativity and its relationship with the development of the intellect is unclear. S.E. Golan has stated, "What is needed for the understanding of the relationship

between creativity and intelligence is not only the data at the correlational level, but conceptual reorganization as well." (cited in Kagan, 1967).

I believe that schools work too diligently at labeling children. Routinely, the labels are incorrect, possibly detrimental, and funnel children down predetermined, often irreversible paths. The attempt to label children as creative or as having creative ability is premature to the debut of the definition for creativity itself. Indeed, the very nature of creativity warrants a cautious pace. One should be hesitant to use creativity as a descriptor until its ambiguity is unveiled and theory as it relates to and manifests itself in education, becomes more clearly defined and measurable. Finally, it is time for researchers to remove the blind-folds and correlate their research so that they might better understand the elephant before them.

References

Anastasi, A. (1968). Psychological testing. New York: Allyn and Bacon.

Baer, J. (1993, December). Why you shouldn't trust creativity tests. Educational Leadership, pp. 80-82.

Crockerberg, S. B. (1972). Creativity tests: a boon or a boondoggle for education? Review of Educational Research, 42, 27-45.

Davis, G. A. (1975). In frumious pursuit of the creative person. The Journal of Creative Behavior, 9, 75-87.

Davis, G. A. & Belcher, T.L. (1971). How shall creativity be measured? Torrance Tests, RAT, Alpha Biographical, and IQ. The Journal of Creative Behavior, 5, 153-161.

Guilford, J. P. (1972). Some misconceptions regarding measurement of creative talents. The Journal of Creative Behavior, 5, 77-87.

Harre, R. & Lamb, R. (Eds.). (1983) The encyclopedic Dictionary of Psychology. Cambridge: The MIT Press.

Hocevar, D. (1979). Measurement of creativity: review and critique. (Report No. TM 009 556). Denver, CO: Rocky Mountain Psychological Association. (ERIC Document Reproduction Service No. ED 175 916)

Kagan, J. (Ed.). (1967). Creativity and learning. Boston: Beacon Press.

Torrance, E.P. & Hall, L.K. (1980). Assessing the further reaches of creative potential. The Journal of Creative Behavior, 1, 1-19.

Treffinger, D. J. & Poggio, J.P. (1972). Needed research on the measurement of creativity. The Journal of Creative Behavior, 6, 253-267.

Simpson, J. A. & Weiner, E.S.C. (Eds.). (1989). The oxford english dictionary (vol. IX). Oxford: Clarendon Press.

Wakefield, J. F. (1991). The outlook for creativity tests. The Journal of Creative Behavior, 25, 184-193.