
Intelligent Testing With Torrance

James C. Kaufman

California State University at San Bernardino

John Baer

Rider University

ABSTRACT: Applying the idea of “intelligent testing” (Kaufman, 1979, 1994) to creativity assessment could broaden our understanding of creativity assessment and would be in accord with the vision of E. Paul Torrance. Such intelligent creativity testing would not seek a single creativity score. Testers would consider patterns of scores in different domains and interpret them for comparative strengths and weaknesses in various areas of creative activity. Specific suggestions on how this might be done are offered.

From 1974 to 1979, two leaders in the field of psychological assessment were both at the University of Georgia: Dr. E. Paul Torrance and Dr. Alan S. Kaufman. Despite their proximity (and personal friendship), they never collaborated on a project together. The goal of this article is to apply Kaufman’s ideas about intelligent testing to the Torrance Tests and Torrance’s work and to suggest how the integration of Torrance and Kaufman might represent a step forward for the field of creativity.

For many years, the point of being assessed with an IQ test was to get one, two, or three scores—Verbal, Performance, and Full-Scale IQs—that were often treated as magic numbers that could open or shut doors of opportunity. If one or more IQs were above a certain point, better school or work opportunities might await. If IQs were below a certain point, someone might be eligible for federal support or assistance.

Creativity tests have no such power. They are sometimes used with gifted programs, but the impact of scoring well on a measure of creativity is quite less than that of scoring well on an IQ test. Yet creativity assessments, like IQ tests, are geared toward global numbers or scores. A popular philosophy of IQ testing, however,

that disdains global scores and has had a tremendous influence on the field is that of “intelligent testing” (Kaufman, 1979, 1994). Using this system, the tester is elevated above the test. The global scores mean little by themselves. The key is interpreting the scores in context. The persons administering the test are expected to use their qualifications and training and to bring their own experience to the testing session. In this manner, the tester can help the child or adult being tested by understanding and interpreting a wide range of behaviors, making inferences about any observed problem-solving strategies, and directly applying the latest theories and research results directly to the person’s specific set of scores. Every aspect of psychology is brought into play to interpret a profile of scores in the context of accumulated research, theory, and clinical practice. This profile is used to help solve problems and create solutions for the person tested—that is, providing answers to the referral questions—not merely as a label or classification system (Kaufman, 1979, 1994).

We believe this approach can be applied to creativity research. A qualified tester would be well versed in the fields of social, cognitive, and educational psychology (among others). The pattern of scores in the different domains could be interpreted for its comparative strengths and weaknesses. Rather than merely producing a single number that is of little use to a student, this new domain-specific creativity could help students dis-

The authors would like to thank Alan S. Kaufman and Scott Barry Kaufman for helpful comments and suggestions on an earlier draft.

Correspondence and requests for reprints should be sent to James C. Kaufman, Learning Research Institute, California State University at San Bernardino, Department of Psychology, 5500 University Parkway, San Bernardino, CA 92407. E-mail: jkaufman@csusb.edu

cover and validate areas of creative talent in themselves. In addition, an administrator using the intelligent testing approach could look for signs of insufficient motivation, a thinking style that might conflict with the task, or other areas that could be improved for enhanced creative potential.

This concept, we believe, would be in line with Torrance's original aims in the development of his divergent thinking tests (the Torrance Tests of Creative Thinking or TTCT; Torrance, 1966, 1974). Torrance did not design his tests for the use to which they are most commonly put these days—identification of students for gifted/talented programs (Kim, 2006). His primary goals in developing these tests were to help us better understand the human mind and its functioning; to find ways to better individualize instruction, including remedial and psychotherapeutic interventions; to evaluate the effectiveness of educational programs; and to become more sensitive to latent potential in people.

With these goals in mind, his first tests provided subscores in the areas of fluency, flexibility, originality, and elaboration (Torrance, 1966, 1974) based on the earlier work on divergent production of Guilford (1956, 1967; Guilford & Hoepfner, 1971). This stimulated a great deal of research, some of which did not support the independence of these four subtests. Hocevar's analyses (1979a, 1979b), for example, suggested that fluency was the only variable of the four actually being measured; similarly, Dixon (1979) showed that originality scores were so dependent on fluency scores as to make them redundant. Factor analytic studies (e.g., Clapham, 1998; Heausler & Thompson, 1988) have also sometimes suggested that the TTCT measures a single factor, and Treffinger (1986) argued that the available evidence did not warrant the use of these four subscale scores as independent measures.

Regardless of any structural or conceptual criticisms of the TTCT, there are many studies that lend support for its validity. Torrance (1972a, 1972b, 1990) reported more than a dozen studies by several researchers showing short- and long-term predictive validity of the tests. These studies included subjects ranging in age from kindergarten through adults. Torrance and Safter (1989), for example, used one of the subtests (Just Suppose) to conduct a long-range study looking at predictive validity and found a solid relationship after 20 years to creative achievement. Plucker (1999) reanalyzed data from a different Torrance longitudinal

study and found that divergent thinking accounted for three times more variance in creative achievement than traditional IQ tests.

The validity of these confirmatory studies has, in turn, been challenged (Baer, 1993; Crockenberg, 1972), just as studies showing no correlation between the tests and later creative accomplishment have been found wanting (Plucker, 1999; Torrance, 1972a). It appears fair to conclude that even if perfect consensus has been elusive regarding the validity of the TTCT, there is no question that they are the most widely used tests of creativity (Baer, 1993; Kim, 2006; Torrance & Presbury, 1984).

In more recent editions of the TTCT, Torrance changed the scoring procedures and introduced 15 new measures: 2 norm-referenced and 13 criterion-referenced (Torrance, 1990, 1998). For example, in addition to the original norm-referenced subtests of fluency, originality, and elaboration (flexibility has been dropped), measures of abstractness of titles and resistance to premature closure have been added to the scoring of the figural test. These new measures open the possibility of more intelligent creativity testing if the various subtests prove to be valid and reliable measures. One component of intelligent testing is analyzing subtests to look for strengths and weakness and patterns of abilities. The newer TTCT also may offer this promise.

We find it interesting—perhaps even prophetic—that the TTCT has for many decades come in two forms, verbal and figural, which suggests that Torrance was long aware of a need to look at how creativity may be different in different domains. The issue of domain specificity has in the past decade become an increasingly important one in creativity theory and research (Baer, 1993, 1998; Kaufman & Baer, 2004, 2005a; Plucker, 1998, 2005; Plucker & Beghetto, 2004). An effort to include both domain-general and domain-specific aspects of creativity is a primary motivation for the theory of creativity that we have been developing, the APT model (Baer & Kaufman, 2005a; Kaufman & Baer, 2005a, 2005b, in press). It is a model that would allow testing of creativity-relevant skills and interests both in different domains and at different levels of domain specificity or generality.

This kind of theory-based testing could provide a rich individual profile of a test taker's creative abilities and interests; in doing so, it would move forward Paul Torrance's (1966, 1974) original goals of helping us

better understand the functioning of the human mind, of finding means to better individualize instruction, of allowing more complete and valid evaluation of program effectiveness, and of discovering latent creative potential in people. It also could provide a better tool for the identification and placement of students in appropriate gifted and talented educational programs.

Because the details of the model are presented elsewhere (Baer & Kaufman, 2005a, 2005b; Kaufman & Baer, 2005b; in press), we will simply summarize the key features of this theory to demonstrate how such a theory might make possible the kind of intelligent testing in the realm of creativity that we have begun to see in intelligence testing. The APT model is based (somewhat whimsically, perhaps, as some of our reviewers have noted) on a large amusement park. (The APT Model = the Amusement Park Theoretical Model.) In an amusement park, there are initial requirements (e.g., a ticket) that apply to all areas of the park. Similarly, there are initial requirements that, to varying degrees, are necessary to creative performance in all domains (e.g., intelligence, motivation).

Amusement parks also have *general thematic areas* (e.g., at Disney World one might select among EPCOT, the Magic Kingdom, the Animal Kingdom, and Disney-MGM Studios), just as there are several different general areas in which someone could be creative (e.g., the arts, science). Once in one type of park, there are sections (e.g., Fantasyland, Tomorrowland), just as there are *domains* of creativity within larger *general thematic areas* (e.g., physics and biology are domains in the *general thematic area* of science). These domains in turn can be subdivided into *microdomains* (e.g., in Fantasyland one might visit Cinderella's Castle or It's a Small World; in the domain of psychology, one might specialize in cognitive psychology or social psychology).

As an example, if one were interested in assessing the creative abilities of a subject in terms of poetry writing creativity, one might start by assessing *initial requirements* such as a certain minimal level of intelligence as well as skill in the *general thematic area* of language. Next one might assess skills in certain *domains* especially relevant to writing poetry (e.g., metaphor-generating ability; see Baer, 1996, for other examples of such domain-specific skills related to creativity in poetry and how such skills can be trained). Finally, if one were interested only in the ability to write haiku (and not other kinds of poetry), one might

evaluate skills in specific microdomains related only to that kind of poetry.

If, however, one were interested in a student's creative potential in the area of physical science, the hierarchy of skills that one would evaluate would be quite different. The *initial requirements* might be similar (e.g., a certain level of intelligence, for example), but skills from very different *general thematic areas* would be of interest (e.g., verbal skills would be less important, and the ability to understand and to generate unusual mathematical ideas would be of much greater interest). The differences would become even greater as one moved down the hierarchy to *domains* and *microdomains*.

Motivation also could be assessed at different levels of such a hierarchy. For example, a student might have strong intrinsic motivation at the level of the *general thematic area* of science, and this would indicate a tendency toward creative productivity in the sciences in general. Another student may have extremely high intrinsic motivation only in the *domain* of marine science, however, which predicts a greater likelihood of creativity in that domain but not in other sciences. Or a student's interest at a given point in time might be even more narrowly focused on a *microdomain* (for example, a student might have great interest in the reproductive success of certain kinds of mollusks in different environments but show little interest in other areas of marine science).

We are currently conducting studies to map this hierarchy of creativity-relevant abilities and interests more precisely. Our longer-term goal is to use this theory to guide the development of assessment devices tied to each element of the model. This will make possible a better understanding of human cognition, intelligence, and individual profiles of creative abilities and interests.

A further goal would be to enable the same types of profile analysis across different creativity domains that are currently being performed across different IQ subtests (after being pioneered by Kaufman, 1979, 1994). Such a method would enable us to more fully understand a student's creativity, much as Kaufman's intelligent testing methods enable us to better understand a student's intelligence. Using a student's strengths and weakness at all levels of our APT model, we can help provide individualized instruction and better identify and place students in appropriate gifted and talented educational programs.

In pursuing these research ideas, we would be following the path-breaking footsteps of E. Paul

Torrance. In addition to being a true pioneer in creativity assessment and testing, Torrance had a unique and far-reaching vision. Even 4 decades ago, he recognized the need for intelligent—not single-scored, one-type-fits-all—creativity testing.

References

- Baer, J. (1993). *Divergent thinking and creativity: A task-specific approach*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Baer, J. (1996). The effects of task-specific divergent-thinking training. *Journal of Creative Behavior*, 30, 183–187.
- Baer, J. (1998). The case for domain specificity in creativity. *Creativity Research Journal*, 11, 173–177.
- Baer, J., & Kaufman, J. C. (2005a). Bridging generality and specificity: The amusement park theoretical (APT) model of creativity. *Roepers Review*, 27, 158–163.
- Baer, J., & Kaufman, J. C. (2005b). Whence creativity? Overlapping and dual aspect skills and traits. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 313–320). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Clapham, M. M. (1998). Structure of the figural forms A and B of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 58, 275–283.
- Crockenberg, S. B. (1972). Creativity tests: A boon or boondoggle for education? *Review of Educational Research*, 42, 27–45.
- Dixon, J. (1979). Quality versus quantity: The need to control for the fluency factor in originality scores on the Torrance Tests. *Journal for the Education of the Gifted*, 2, 70–79.
- Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, 53, 267–293.
- 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.
- Hocevar, D. (1979a). Ideational fluency as a confounding factor in the measurement of originality. *Journal of Educational Psychology*, 71, 191–196.
- Hocevar, D. (1979b). The unidimensional nature of creative thinking in fifth-grade children. *Child Study Journal*, 9, 273–278.
- Heausler, N. L., & Thompson, B. (1988). Structure of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 48, 463–468.
- Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.
- Kaufman, J. C., & Baer, J. (2004). Hawking's haiku, Madonna's math: Why it's hard to be creative in every room of the house. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 3–19). Washington, DC: American Psychological Association.
- Kaufman, J. C., & Baer, J. (Eds.). (2005a). *Creativity across domains: Faces of the muse*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kaufman, J. C., & Baer, J. (2005b). The amusement park theory of creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 321–328). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kaufman, J. C., & Baer, J. (in press). Creativity theory for gifted education: The Amusement Park Theoretical (APT) model. *Sobredotação*.
- Kim, K. H. (2006). Can we trust creativity tests: A review of the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 18, 3–14.
- Plucker, J. A. (1998). Beware of simple conclusions: The case for the content generality of creativity. *Creativity Research Journal*, 11, 179–182.
- Plucker, J. A. (1999). Is the proof in the pudding? Reanalyses of Torrance's (1985 to present) longitudinal data. *Creativity Research Journal*, 12, 103–114.
- Plucker, J. A. (2005). The (relatively) generalist view of creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 307–312). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Plucker, J. A., & Beghetto, R. A. (2004). Why creativity is domain general, why it looks domain specific, and why the distinction does not matter. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 153–167). Washington, DC: American Psychological Association.
- Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking: Norms-technical manual*. Lexington, MA: Personal Press.
- Torrance, E. P. (1972a). Career patterns and peak creative achievements of creative high school students twelve years later. *Gifted Child Quarterly*, 16, 75–88.
- Torrance, E. P. (1972b). Predictive validity of the Torrance Tests of Creative Thinking. *Journal of Creative Behavior*, 6, 236–252.
- Torrance, E. P. (1974). *The Torrance Tests of Creative Thinking: Norms-technical manual*. Princeton, NJ: Personal Press.
- Torrance, E. P. (1990). *The Torrance Tests of Creative Thinking: Norms-technical manual*. Bensenville, IL: Scholastic Testing Service.
- Torrance, E. P. (1998). *The Torrance Tests of Creative Thinking: Norms-technical manual*. Bensenville, IL: Scholastic Testing Service.
- 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.
- Torrance, E. P., & Safter, H. T. (1989). The long range predictive validity of the Just Suppose Test. *Journal of Creative Behavior*, 23, 219–223.
- Treffinger, D. J. (1986). Research on creativity. *Gifted Child Quarterly*, 30, 15–19.