

Article

Why Creativity Isn't in IQ Tests, Why it Matters, and Why it Won't Change Anytime Soon Probably

James C. Kaufman

Department of Educational Psychology, Neag School of Education, University of Connecticut, 2131 Hillside Road, Storrs, CT 06269-3007, USA; E-Mail: james.kaufman@uconn.edu

Academic Editor: Robert J. Sternberg

Received: 18 June 2015 / Accepted: 28 July 2015 / Published: 7 August 2015

Abstract: Creativity is a part of most theories of intelligence—sometimes a small part and sometimes a large part. Yet even IQ tests that assess aspects of intelligence that supposedly reflect creative abilities do not actually measure creativity. Recent work has argued that intelligence and creativity are more conceptually related than we have thought. In addition, creativity offers a potential way to counter issues of test bias from several different angles. That said, inherent difficulties in measuring creativity and inherent sluggishness in the test industry mean the odds are small that creativity will find its way into IQ tests as currently defined. However, there remain other potential possibilities in related fields.

Keywords: IQ testing; creativity; intelligence; creativity assessment

1. Introduction

There are many challenges facing IQ testing, and how the field responds will likely determine the extent of their continued use in the future. A small sampling of issues includes how tests can incorporate technological advances (indeed, most intelligence tests are oddly stuck in the hard-copy era), how test developers can respond to such challenges as the RTI (Response to Intervention) movement, and how trainers can encourage intelligent testing [1]. There are broader concerns, however, which are rooted in some of the same basic questions that have been asked for the last century. Do IQ tests actually measure intelligence? How well do they predict real-world success? Do IQ tests reflect current beliefs about the vast expanse of intellectual abilities?

In some ways, it is unfair to expect one test—whether we mean an IQ test, an academic achievement test, or an admissions test such as the SATs or GREs—to account for everything. We wouldn't expect one blood test to yield a diagnosis for all possible diseases. Yet in common perception and often in practice, we allow a handful of scores to determine virtually everything about a person.

In an ideal world, every child being assessed would be given a complete battery of cognitive and non-cognitive tests, just as any college applicant would be evaluated on a portfolio of past work, scores, recommendations, interviews, and essays. Both IQ tests and college admission tests are often defended by saying they are only intended to be used as part of a complete work-up; I've called this argument the Lucky Charms paradox [2]. In commercials, a child is portrayed eating Lucky Charms as part of a nutritious breakfast, complete with orange juice, scrambled eggs, yogurt, and toast. When seen this way, Lucky Charms is a good thing. But in real life, mornings are frenetic and filled with oversleeping kids, hurried showers, and a quick bowl or two of Lucky Charms gulped down before the school bus. Despite the best intentions, Lucky Charms become the only source of breakfast nutrition.

Moving from breakfast cereal to tests, the SATs are often the only numbers given weight for most college admissions. Low GRE scores can keep students with otherwise exceptional records out of top graduate schools. And IQ tests, all by themselves, hold a tremendous amount of power. It is common for students to be initially recommended for gifted programs but then denied if their IQ score is one point lower than the cut-off [3,4].

IQ can take on a literally life-or-death role given that an IQ of 70 is the cutoff for whether a prisoner can be executed [5]. Even though technically a classification of Intellectual Disability (Mental Retardation) requires both an IQ and a measure of adaptive behavior (such as the Vineland—II) to be below 70, in reality it is the IQ test that bears the burden of determining a criminal's fate [6]. Further, sometimes a live-or-die verdict rests on less than 1 IQ point. IQ test norms get out of date at the rate of 0.3 points per year, and many researchers and clinicians argue that a person's IQ must be adjusted by 0.3 points for each year the norms are out of date (Flynn, 2009)—even if that adjustment produces an IQ of 69.4 or 70.7. As Cecil Reynolds and his colleagues argue: “The importance of understanding and assessing mental retardation in criminal defendants has become critical, indeed a true matter of life and death, in capital felony cases. No one's life should depend on when an IQ test was normed” [7]. Even more pertinent, in this high-tech age brimming with sophisticated theories of intelligence, creativity, and non-cognitive variables, how can key decisions like gifted placements or capital punishment verdicts rest almost exclusively on *g*, and be able to be overturned by a point or two?

With power comes responsibility. What aspects of intelligence are not measured by intelligence tests? Here there is natural divergence between IQ test developers and intelligence theorists, yet at least one construct is present in all theories but still omitted from any IQ test: Creativity.

2. Creativity in IQ-Related Theories and IQ Tests

Most current IQ tests use the Cattell-Horn-Carroll (CHC) model [1] either explicitly or implicitly within their theoretical frame work [8]. Given that CHC is the theory most commonly used in IQ tests and is also the most frequently used non-*g* perspective used in creativity-intelligence studies, I will primarily focus on this theory.

The CHC model is an outgrowth of the Cattell-Horn Gf-Gc theory of fluid and crystallized intelligence [9] and Carroll's (1993) Three-Stratum Theory [10]. Much of current CHC theory is most rooted in Horn's expanded model (e.g., [11]). Although Gf would seem to be a natural match with creativity abilities, with its emphasis on novel problem-solving, creativity is placed as part of Glr (Long-term storage and retrieval) in the current CHC model [12,13].

Someone high in Glr is able to store information in long-term memory and then retrieve it when needed [14]. Glr is used when you learn the name of your daughter's fifth-grade teacher and then again when you remember the name of your own fifth-grade teacher. You are still using Glr when you help your daughter with her homework and can recall your own long-ago knowledge of the state capitals and then link this memory of Sacramento being the capital of California with last night's news story about the lottery winner living in Sacramento.

Glr has two distinct components: learning efficiency (how well you can both learn and retain new information) and fluency (the ability to rapidly recall many things). There are many "narrow" abilities underneath these two components; one such ability is Idea Production, which can include five types of fluency (associational, expressional, ideational, word, and figural), figural flexibility, and originality [15]. Fluency, which is this usage refers to the ability to generate many different ideas; flexibility, or the ability to generate many different categories of ideas; and originality, which is generating particularly rare and unusual ideas, are also key components in creativity. They represent a large portion of Guilford's Divergent Production [16] within his Structure of Intellect model.

Although Glr is measured in many current IQ tests, the focus is rarely on fluency (and never on flexibility or originality). The Woodcock-Johnson IV Tests of Oral Language [17] include Retrieval Fluency. The Woodcock-Johnson IV Tests of Achievement (WJ-IV; [18]) include three Fluency subtests (Sentence Reading, Math Facts, and Science Writing) in the core battery and Word Reading Fluency in the extended battery. The Kaufman Test of Educational Achievement-3 (KTEA-3; [19]) includes nine Fluency subtests: Word Recognition, Silent Reading, Decoding, Math, Writing, Associational, Reading, Oral, and Academic. The NEPSY-II [20] includes Design Fluency and Word Generation (which was originally called Verbal Fluency). Yet most of these subtests are measuring the broader type of fluency, which is less related to creativity. As J.C. Kaufman *et al.* [15] discuss, the only four subtests that measure fluency in a creativity-relevant way are the WJ IV's Retrieval Fluency, the KTEA-3's Associational Fluency, and the two NEPSY-II subtests. Although recent studies have found that Glr does indeed predict creativity [21,22], the construct as measured by IQ tests is far removed from conceptions of creativity.

The CHC model is certainly not the only (or necessarily best) theory of intelligence. The PASS model (Planning, Attention, Simultaneous, and Successive) [23], a cognitive processing approach rooted in the neuropsychological work of Luria [24], is also often used in IQ tests. Creativity likely lies in the Planning component [25]. Tests that use the PASS model do measure Planning, but such assessments have not included anything explicitly related to creativity. In addition, there are several theories that are not (yet) represented in IQ tests that have much to offer.

3. Creativity in Other Theories of Intelligence

Guilford's [16] Structure of Intellect model is not as popular as it once was, but it represented a pioneering step forward. The model introduced the concepts of both divergent and convergent thinking, which are parts of many modern theories of creativity (e.g., [26]).

A modern theory of intelligence that emphasizes creativity is Sternberg's [27–29] theory of successful intelligence. An earlier version, the Triarchic Theory of Intelligence [30], proposed three intelligences: Analytic, Practical, and Creative. The revised theory comprises three subtheories. Most related to creativity is the experiential subtheory, which focuses on how people adjust to novelty and automatize information processing. Although many measures of Sternberg's theories have been created and used for college admissions [31], they have not yet been incorporated into IQ tests (see, e.g., [32]).

Gardner's well-known theory of multiple intelligences [33,34] (encompassing interpersonal, intrapersonal, spatial, naturalistic, language, logical-mathematical, bodily-kinesthetic, and musical) includes a wide range of abilities that would require differing amounts of creativity. Further, Gardner's *Creating Minds* [35] contains case studies of eminent creative individuals who embody his intelligences. Although Gardner's work has had extensive influence in the schools, his ideas have also not yet been incorporated into IQ tests.

4. The Relationship between Creativity and Intelligence

There have been hundreds of papers devoted to this topic, and given that my emphasis is on creativity as part of intelligence (as opposed to creativity representing a different construct from intelligence), I will only briefly review this literature. Most studies on the topic tend to assume a generalist perspective on both intelligence (typically using group-based tests of *g*) and creativity (typically using divergent thinking tests). Creativity and intelligence, under these circumstances, tend to correlate at a small but significant level [36,37].

Many argue that the two constructs are more closely related than such studies would indicate. Silvia [38,39] suggests the relationship is underestimated because the studies are limited by only looking at observable scores (*i.e.*, performance on an intelligence test). Jung [40] notes that intelligence can be seen as problem solving at an everyday level (e.g., [41]), whereas creativity may represent problem solving for less common issues (e.g., [42]). Others argue that creativity and intelligence are both cognitive functions [43] or that divergent thinking is simply an executive cognitive function [44].

Much of the more recent work on creativity and intelligence has focused more on *Gf* instead of *g* or *Gc*. Benedek, *et al.* [45] studied underlying executive functions behind both divergent thinking and *Gf*. They found that the ability to notice small changes (called updating; [46]) predicted both *Gf* and divergent thinking; the ability to stifle a natural response (or inhibition) also predicted divergent thinking. Other studies have used metaphor creation as a creativity measure instead of divergent thinking and a spectrum of CHC components instead of just *g* and have found much higher relationships between creativity and intelligence than past studies (*i.e.*, [47,48]).

A pair of studies by Silvia and colleagues indicated that the relationship between *Gf* and creativity is mediated by other cognitive mechanisms. Nusbaum and Silvia [49] found that *Gf* predicted creativity but also predicted how well people could use a more efficient strategy to improve their scores. Beaty

and Silvia [50] studied divergent thinking over time. Participants in the sample with higher Gf produced creative initial ideas but slowed over time; in comparison, those participants with lower Gf showed much more improvement in their idea generation when given more time. Looking broadly at these studies suggests that the basic approach to seeing how creativity and intelligence relate to each other may not be enough. Creativity and intelligence show an intricate relationship with many cognitive and situational mediators. It may be instinctually appealing to consider them as completely different concepts, but some have argued that this separation can hurt children who are being assessed [51].

5. Why is Creativity Important? The Issue of Bias

Even if one readily accepts that creativity is a part of intelligence and that it is not satisfactorily measured on IQ tests, there is the larger question of whether creativity's absence is cause for concern. It is easy to list the many positive outcomes connected with creativity, but most components of intelligence would have their own list of beneficial associations. As I have argued elsewhere [52,53], creativity has particular implications for fairness and non-biased assessment.

IQ tests have been criticized in two different ways for being biased. A common layperson criticism approach is to note the significant differences that occur between males and females (although such differences are much more present in standardized achievement tests) and among ethnic groups. Some researchers argue that these measures reflect actual differences [54]. Others point to the discrepancy between socioeconomic status and opportunities across ethnicities [27], whereas still others argue that current ability measures do not incorporate enough aspects of intelligence to truly reflect a person's "global" ability [29].

Another proposed explanation is stereotype threat. A multitude of studies (see, e.g., [55]) have found that individuals feel stress when placed in a situation where they run the risk of confirming a negative stereotype about their group (e.g., ethnicity). This stress can then cause poor performance [56,57]. People can be aware of stereotypes about intelligence when they do not endorse such views; indeed, even people who are the targets of such stereotypes are cognizant of these views. An example of a manifestation might be if an African American test-taker is aware of these stereotypes and is worried about confirming such negative views and, as a result, feels added stress. Such added stress may increase cognitive load and reduce working memory [58], thereby leading to lowered test performance.

Beyond simply focusing on group differences, there is a psychometric approach to bias that offers a more nuanced approach. In this method, the question is whether a test might be measuring different things in different groups. For example, a test may measure vocabulary in Caucasians, but measure something else (such as exposure to American culture) in a Hispanic American population [59,60].

If a test measures something different from its intended purpose for specific groups, then it can be considered to be biased against such groups. A measure can be considered to be fair to the extent any resultant score only reflects (a) variables associated with the construct being measured, and (b) random variance from error. If any systematic error occurs because of a person's group membership and, thus, that person receives a lower score, then the test is biased [61].

Creativity may be a reason why some groups systematically get items incorrect. Several scholars [62–64] have suggested that Caucasians approach some tasks (such as hearing a story and then recounting

specific details) in the way that the test creators intended. In contrast, African Americans may emphasize the narrative in the story and answer more creatively. Such answers would then be marked incorrect.

Creativity also may help with the issue of test bias in other ways. If stereotype threat is a reason why some groups receive lower IQs, it is worth noting that creativity tests and rated work tend to show no differences across ethnicity [65–67]. However, African Americans see themselves as more creative than Caucasians in several domains [68]. In another study, African Americans and Caucasians were asked to provide self-ratings on both intelligence and creativity [69]. There was an interaction with education; among less-educated individuals, ethnic differences were notably higher for intelligence than creativity. Among the higher-educated individuals, African Americans rated themselves higher for creativity than did Caucasians; the opposite effect was found for intelligence. The result of these findings may be that given the higher self-concept that African Americans have for creativity, an IQ test that includes creativity may reduce stereotype threat. Further, there is a strong chance that the ethnicity gap noted in IQ tests would not be found in items emphasizing creativity.

6. Creativity and IQ Tests: Possibilities, Realities, and Ironies

Although the field of creativity has been around for more than 65 years and has been especially fertile over the last two decades, measurement is still a tricky subject. The most common assessment remains divergent thinking, much as it was in the 1950s. Researchers often use either the Torrance Tests of Creative Thinking (TTCT; [70]) or a comparable measure based on Guilford's [16] ideas [71]. Such tests are traditionally scored for fluency, flexibility, and originality (all abilities that would theoretically comprise G_{lr}).

Could a subtest that measures divergent thinking ("Creative Fluency") be added to an IQ test? Instead of asking test-takers to name many different examples from a category (such as types of flowers), could they be asked more open-ended questions such as uses for a pencil or what might happen if people could fly? Certainly, it would be possible to score such responses for fluency; just as IQ test examiners are highly trained as to what constitutes a correct answer, so too are TTCT scorers trained to reliably determine if a response is relevant (and thus counts for fluency) or irrelevant [72].

Even scoring for originality would be possible. The TTCT has a matrix of responses and points are assigned based on the rarity of response. In research studies, all responses are pooled and originality is scored based on the frequency a particular answer is given. Such matrices would be easy to compile during test standardization; one such matrix was applied to determine the originality of responses on the Verbal Fluency subtest of the McCarthy Scales of Children's Abilities [73]. Just as current IQ tests have responses that carry different weights based on "correctness", so too could a scoring system be implemented that assigns increased points for more rare responses. Such an assessment could also be translated on-line without notable changes to the process [74].

There are criticisms of divergent thinking, of course. Even its supporters agree that it is only one aspect of creativity [75]. The broad verbal-figural categories are not dissimilar to the early atheoretical Verbal-Performance structure of past Wechsler scales (e.g., [76]). The validity of such measures has been both challenged [77] and defended [78]. Adding divergent thinking to IQ tests would probably be the easiest possibility, but not necessarily the optimal choice. Indeed, consider that one common criticism

of IQ tests is that they have largely remained stagnant over the last century [32]. The selection of a creativity measure primarily rooted in work more than six decades old may not be a step forward.

Another popular measure of creativity is the Consensual Assessment Technique [79,80], in which actual creative products are rated by content experts. Such experts tend to agree at strikingly high rates, even if they represent different areas of expertise [81,82] or a different amount of expertise [83,84]. Only pure novices do not show agreement with experts [85,86].

This method would be impractical on a large scale for reasons of time and cost, but an important take-away is that an expert's opinion carries a certain amount of weight. Just as an examiner need not be an expert in motivation or personality to make valuable notes about a child's behavior during the test, particular care can be made to note responses that are notably original, aesthetically pleasing, humorous, or clever. Such behavioral observations are consistent with a teacher using slight variations from a typical lesson plan to allow student creativity to develop [87].

Even the Consensual Assessment Technique could eventually be utilized on a larger scale. Computers can be trained to grade essays – not to mention to play chess and develop unique recipes [72]. There is no reason to assume that given enough programming, computers could not be trained to mimic human raters in evaluating creative work (at least creative writing). How might computers be trained? Consider the many textual analysis programs (e.g., WordNet, Linguistic Integrative Word Count, and Latent Semantic Analysis) that use advanced computational modeling techniques to assess semantic patterns and associations. Many of these models are based on self-learning algorithms, meaning that there is minimal training time and that the ability to detect whatever construct is being examined improves with experience. Some of the programs (such as Latent Semantic Analysis) are already being used to automatically score essays based on human feedback; this same basic principle can be applied to creative writing. Just as current systems have trained computer programs to mimic experts in grading essays for writing quality, grammar, and sentence structure, so too can programs be trained to detect creativity. Infinite facets of a writing sample can be detected (across many possible dimensions, from adjective placement to language choice) and then associated with multiple ratings by experts of the creative aspects of the product. The program could then learn which features are associated with creativity, and to what extent. Certainly, the ability to calculate distances between words or concepts should allow for sophisticated originality scores in divergent thinking-type measures.

I have used creative writing as a main example because I think we are the closest to being able to best measure that domain on a large scale. However, the Consensual Assessment Technique has been used in many others domains, from taking photographs [88] to music compositions [89] to deriving mathematical equations [90] to dramatic performance [91] to answering science questions [92] to cooking [93] to everyday problem-solving [94]. In addition, ratings of different creative products within the same domain tend to be consistent by person [80,95].

Ideally, a multitude of possible domains could be measured to provide a broad perspective on a person's creativity. Although it would be simpler to use a domain-general measure of creativity, there has been a strong movement toward domain-specificity [85,96,97].

Unfortunately, these types of initiatives take money and the willingness to take risks. Test publishers may have the former, but most have shown no indication of the latter. On some level, this is understandable; IQ tests are a business, like making dental floss, and the people making the decisions—even those with backgrounds in psychology or education—do not seem to have the burning passion to advance

the field. It is easy to criticize other people for not spending money, taking chances, or investing in the future. Yet, ironically, what is preventing creativity from being included in IQ tests is a lack of creativity on the part of publishers.

7. The Future of Creativity -and IQ -Assessment?

No matter how much some groups want to keep the status quo, progress moves forward regardless. Video games continue to grow in popularity, both as an activity and a business. There has been a growing interest in the relationship between video games and creativity [98]. Some studies show a link between playing video games and being creative [99,100]. But these are only the start.

Valerie Shute and her colleagues [101–103] have begun what they call “stealth assessment” in video games. They have created a game called Physics Playground, in which the player draws objects on the screen. The drawings then become animated and interact with other objects in the game; as the game progresses, there are a series of puzzles and problems that can be solved by drawing different objects. Shute and her colleagues can quietly test how well the players learn and understand physics—as well as their creativity. Levels can be solved in many different ways, and a player can attempt multiple solutions. It is thus possible to generate scores for fluency, flexibility, and originality (as well as related constructs, such as humor or aesthetics; [101]).

Shute and her colleagues’ work is still at a starting point. There are years of work in continued development, validation, and expansion. The possibilities discussed earlier—especially the use of computerized scoring to make large scale Consensual Assessment Technique feasible—are also likely far away from being a reality. Unfortunately, we are many years removed from a time where researchers such as David Wechsler (or my parents, Alan and Nadeen Kaufman) drove the field. The publishers are the ones with all of the power, and the dominating company, Pearson, is only now slowly translating its existing instruments into computerized format with Q-interactive (see, e.g., [104]).

It used to be that the costs of creating and standardizing a test were so prohibitive that a huge publisher was a necessity. Things can change. There are some aspects of testing that need a face-to-face component [1], but on-line assessment will nonetheless be suitable for many purposes. Collaboration with programmers and using MTurk to obtain a large standardization sample may enable academics or independent companies to tackle the type of new approaches—creative approaches, and ones that also measure creativity—that current industry leaders are unwilling to consider.

Acknowledgments

The author would like to thank Alan S. Kaufman for his suggestions on an earlier draft.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Kaufman, A.S. *IQ Testing 101*; Springer Publishing Company: New York, NY, USA, 2009.

2. Kaufman, J.C.; Kaufman, A.S. It can be very tempting to throw out the baby with the bathwater: A father-and-son commentary on Does IQ really predict job performance? *Appl. Dev. Sci.* **2015**, *19*, doi:10.1080/10888691.2015.1008922, in press.
3. Pfeiffer, S.I. (Ed.). *Handbook of Giftedness in Children. Psychoeducational Theory, Research and Best Practices*; Springer Science + Business Media: New York, NY, USA, 2008.
4. McClain, M.C.; Pfeiffer, S. Identification of gifted students in the United States today: A look at state definitions, policies, and practices. *J. Appl. Sch. Psychol.* **2012**, *28*, 59–88.
5. Kelly, B.; Resnick, P. A single IQ score over 70 supports finding of no intellectual disability, despite conflicting test results and expert testimony. *J. Am. Acad. Psychiatry Law* **2014**, *42*, 515–517.
6. Kaufman, A.S. “In What Way Are Apples and Oranges Alike?” A Critique of Flynn’s Interpretation of the Flynn Effect. *J. Psychoeduc. Assess.* **2010**, *28*, 382–398.
7. Reynolds, C.R.; Niland, J.; Wright, J.E.; Rosenn, M. Failure to apply the Flynn correction in death penalty litigation: Standard practice of today maybe, but certainly malpractice of tomorrow. *J. Psychoeduc. Assess.* **2010**, *28*, 477–481.
8. Keith, T.Z.; Reynolds, M.R. Cattell-Horn-Carroll abilities and cognitive tests: What we’ve learned from 20 years of research. *Psychol. Sch.* **2010**, *47*, 635–650.
9. Horn, J.L.; Cattell, R.B. Refinement and test of the theory of fluid and crystallized intelligence. *J. Educ. Psychol.* **1966**, *57*, 253–270.
10. Carroll, J.B. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*; Cambridge University Press: New York, NY, USA, 1993.
11. Horn, J.L.; Blankson, N. Foundations for better understanding of cognitive abilities. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, 2nd ed.; Flanagan, D.P., Harrison, P.L., Eds.; Guilford: New York, NY, USA, 2005; pp. 41–68.
12. Schneider, W.J.; McGrew, K.S. The Cattell-Horn-Carroll model of intelligence. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, 3rd ed.; Flanagan, D.P., Harrison, P.L., Eds.; Guilford Press: New York, NY, USA, 2012; pp. 99–144.
13. Schneider, W.J. Flanagan, D.P. The relationship between theories of intelligence and intelligence tests. In *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and current Concepts*; Goldstein, S., Princiotta, D., Naglieri, J.A., Eds.; Springer: New York, NY, USA, 2015; pp. 317–340.
14. Flanagan, D.P.; Ortiz, S.O.; Alfonso, V.C. *Essentials of Cross Battery Assessment*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2013.
15. Kaufman, J.C.; Kaufman, S.B.; Lichtenberger, E.O. Finding creativity on intelligence tests via divergent production. *Can. J. Sch. Psychol.* **2011**, *26*, 83–106.
16. Guilford, J.P. *The Nature of Human Intelligence*; McGraw-Hill: New York, NY, USA, 1967.
17. Schrank, F.A.; Mather, N.; McGrew, K.S. *Woodcock-Johnson IV Tests of Oral Language*; Riverside: Rolling Meadows, IL, USA, 2014.
18. Schrank, F.A.; Mather, N.; McGrew, K.S. *Woodcock-Johnson IV Tests of Achievement*; Riverside: Rolling Meadows, IL, USA, 2014.
19. Kaufman, A.S.; Kaufman, N.L. *Kaufman Test of Educational Achievement*, 3rd ed.; NCS Pearson: Bloomington, MN, USA, 2014.

20. Korkman, M.; Kirk, U.; Kemp, S. *NEPSY*, 2nd ed.; Psychological Corporation: San Antonio, TX, USA, 2007.
21. Avitia, M.J.; Kaufman, J.C. Beyond g and c: The relationship of rated creativity to long-term storage and retrieval (glr). *Psychol. Aesthet. Creativity Arts* **2014**, *8*, 293–302.
22. Silvia, P.J.; Beaty, R.E.; Nusbaum, E.C. Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence* **2013**, *41*, 328–340.
23. Das, J.P.; Naglieri, J.A.; Kirby, J.R. *Assessment of Cognitive Processes: The PASS Theory of Intelligence*; Allyn & Bacon: Boston, MA, USA, 1994.
24. Luria, A.R. The functional organization of the brain. *Sci. Am.* **1970**, *222*, 66–78.
25. Naglieri, J.A.; Kaufman, J.C. Understanding intelligence, giftedness, and creativity using PASS theory. *Roeper Rev.* **2001**, *23*, 151–156.
26. Finke, R.A.; Ward, T.B.; Smith, S.M. *Creative Cognition: Theory, Research, and Applications*. MIT Press: Cambridge, MA, USA, 1992.
27. Sternberg, R.J. *Successful Intelligence*; Plume: New York, NY, USA, 1997.
28. Sternberg, R.J. The theory of successful intelligence. *Rev. General Psychol.* **1999**, *3*, 292–316.
29. Sternberg, R.J.; Kaufman, J.C.; Grigorenko, E.L. *Applied Intelligence*; Cambridge University Press: Cambridge, UK, 2008.
30. Sternberg, R.J. *Beyond IQ: A Triarchic Theory of Human Intelligence*; Cambridge University Press: New York, NY, USA, 1985.
31. Sternberg, R.J.; Coffin, L.A. Admitting and developing “new leaders for a changing world”. *N. Engl. J. High. Educ.* **2010**, *24*, 12–13.
32. Sternberg, R.J.; Kaufman, J.C. Innovation and intelligence tests: The curious case of the dog that didn’t bark. *Eur. J. Psychol. Assess.* **1996**, *12*, 167–174.
33. Gardner, H. *Intelligence Reframed: Multiple Intelligences for the 21st Century*; Basic Books: New York, NY, USA, 1999.
34. Gardner, H. *Multiple Intelligences: New Horizons*; Perseus: New York, NY, USA, 2006.
35. Gardner, H. *Creating Minds*; Basic Books: New York, NY, USA, 1993.
36. Barron, F.; Harrington, D.M. Creativity, intelligence, and personality. *Annu. Rev. Psychol.* **1981**, *32*, 439–476.
37. Kim, K.H. Can only intelligent people be creative? *J. Second. Gift. Educ.* **2005**, *16*, 57–66.
38. Silvia, P.J. Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personal. Individual Differ.* **2008**, *44*, 1012–1021.
39. Silvia, P.J. Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan 1965. *Creativity Res. J.* **2008**, *20*, 34–39.
40. Jung, R.E. Evolution, creativity, intelligence, and madness: “Here Be Dragons”. *Front. Psychol.* **2014**, *5*, 784, doi:10.3389/fpsyg.2014.00784.
41. Kaufman, S.B.; DeYoung, C.G.; Reis, D.L.; Gray, J.R. General intelligence predicts reasoning ability even for evolutionarily familiar content. *Intelligence* **2011**, *39*, 311–322.
42. Simonton, D.K. Creativity and discovery as blind variation: Campbell’s (1960) BVSR model after the half-century mark. *Rev. General Psychol.* **2011**, *15*, 158–174.

43. Silvia, P.J. Intelligence and creativity are pretty similar after all. *Educ. Psychol. Rev.* **2015**, 1–8, doi:10.1007/s10648-015-9299-1.
44. Gilhooly, K.J.; Fioratou, E.; Anthony, S.H.; Wynn, V. Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *Br. J. Psychol.* **2007**, *98*, 611–625.
45. Benedek, M.; Jauk, E.; Sommer, M.; Arendasy, M.; Neubauer, A.C. Intelligence, creativity, and cognitive control: the common and differential involvement of executive functions in intelligence and creativity. *Intelligence* **2014**, *46*, 73–83.
46. Miyake, A.; Friedman, N.P. The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Curr. Dir. Psychol. Sci.* **2012**, *21*, 8–14.
47. Beaty, R.E.; Silvia, P.J. Metaphorically speaking: Cognitive abilities and the production of figurative language. *Memory Cogn.* **2013**, *41*, 255–267.
48. Silvia, P.J.; Beaty, R.E. Making creative metaphors: the importance of fluid intelligence for creative thought. *Intelligence* **2012**, *40*, 343–351.
49. Nusbaum, E.C.; Silvia, P.J. Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence* **2011**, *39*, 36–45.
50. Beaty, R.E.; Silvia, P.J. Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychol. Aesthet. Creativity Arts* **2012**, *6*, 309–319.
51. Kaufman, S.B. *Ungifted: Intelligence Redefined*; Basic Books: New York, NY, USA, 2013.
52. Kaufman, J.C. Non-Biased Assessment: A Supplemental Approach. In *Comprehensive Handbook of Multicultural School Psychology*; Frisby, C.L., Reynolds, C.R., Eds.; Wiley: Hoboken, NJ, USA, 2005.
53. Kaufman, J.C. Using creativity to reduce ethnic bias in college admissions. *Rev. General Psychol.* **2010**, *14*, 189–203.
54. Jensen, A.R. *The g Factor: The Science of Mental Ability*; Praeger: Westport, CT, USA, 1998.
55. Nguyen, H.D. Ryan, A.M. Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *J. Appl. Psychol.* **2008**, *93*, 1314–1334.
56. Steele, C. A threat in the air: How stereotypes shape intellectual identity and performance. *Am. Psychol.* **1997**, *52*, 613–629.
57. Steele, C.; Aronson, J. Stereotype threat and the intellectual performance of African Americans. *J. Personal. Soc. Psychol.* **1995**, *69*, 797–811.
58. Schmader, T.; Johns, M. Converging evidence that stereotype threat reduces working memory capacity. *J. Personal. Soc. Psychol.* **2003**, *85*, 440–452.
59. Reynolds, C.R. Methods for detecting and evaluating cultural bias in neuropsychological tests. In *Handbook of Cross-Cultural Neuropsychology*; Fletcher-Janzen, E., Strickland, T.L., Reynolds, C.R., Eds.; Springer: New York, NY, USA, 2000; pp. 249–285.
60. Reynolds, C.R.; Lowe, P.A.; Saenz, A.L. The problem of bias in psychological assessment. In *The Handbook of School Psychology*; Reynolds, C.R., Gutkin, T.B., Eds.; Wiley: New York, NY, USA, 1999; pp. 549–596.
61. Mackintosh, N. *IQ and Human Intelligence*; Oxford University Press: Oxford, UK, 1998.
62. Heath, S.B. *Ways with Words*; Cambridge University Press: Cambridge, UK, 1983.
63. Kaufman, J.C. *Creativity 101*; Springer: New York, NY, USA, 2009.

64. Manly, J.J.; Miller, S.W.; Heaton, R.K.; Byrd, D.; Reilly, J.; Velasquez, R.J.; Saccuzzo, D.P.; Grant, I.; the HIV Neurobehavioral Research Center (HNRC) Group. The effect of Black acculturation on neuropsychological test performance in normal and HIV positive individuals. *J. Intl. Neuropsychol. Soc.* **1998**, *4*, 291–302.
65. Glover, J.A. Comparative levels of creative ability in Black and White college students. *J. Gen. Psychol.* **1976**, *128*, 95–99.
66. Glover, J.A. Comparative levels of creative ability among elementary school children. *J. Gen. Psychol.* **1976**, *129*, 131–135.
67. Kaufman, J.C.; Baer, J.; Gentile, C.A. Differences in gender and ethnicity as measured by ratings of three writing tasks. *J. Creative Behav.* **2004**, *38*, 56–69.
68. Kaufman, J.C. Self-reported differences in creativity by gender and ethnicity. *J. Appl. Cogn. Psychol.* **2006**, *20*, 1065–1082.
69. Ivcevic, Z.; Kaufman, J.C. The can and cannot do attitude: How self estimates of ability vary across ethnic and socioeconomic groups. *Learn. Individual Differ.* **2013**, *27*, 144–148.
70. Torrance, E.P. *The Torrance Tests of Creative Thinking Norms-Technical Manual Figural (Streamlined) Forms A & B*; Scholastic Testing Service: Bensenville, IL, USA, 2008.
71. Kaufman, J.C.; Plucker, J.A.; Baer, J. *Essentials of Creativity Assessment*; Wiley: New York, NY, USA, 2008.
72. Kaufman, J.C. *Creativity 101*, 2nd ed.; Springer: New York, NY, USA, in press.
73. Kaufman, N.L.; Kaufman, A.S. Creativity in children with minimal brain dysfunction. *J. Creative Behav.* **1980**, *14*, 73, doi:10.1002/j.2162-6057.1980.tb00228.x.
74. Hass, R.W. Feasibility of online divergent thinking assessment. *Comput. Hum. Behav.* **2015**, *46*, 85–93.
75. Torrance, E.P.; Cramond, B. Needs of creativity programs, training, and research in the schools of the future. *Res. Sch.* **2002**, *9*, 5–14.
76. Wechsler, D. *Manual for the Wechsler Intelligence Scale for Children—Revised (WISC-R)*; The Psychological Corporation: New York, NY, USA, 1974.
77. Baer, J. How divergent thinking tests mislead us: Are the Torrance Tests still relevant in the 21st century? The Division 10 debate. *Psychol. Aesthet. Creativity Arts* **2011**, *5*, 309–313.
78. Kim, K.H. The APA 2009 Division 10 debate: Are the Torrance Tests of Creative Thinking still relevant in the 21st century? *Psychol. Aesthet. Creativity Arts* **2011**, *5*, 302–308.
79. Amabile, T.M. *Creativity in Context*; Westview Press: Boulder, CO, USA, 1996.
80. Kaufman, J.C.; Baer, J. Beyond new and appropriate: Who decides what is creative? *Creativity Res. J.* **2012**, *24*, 83–91.
81. Kaufman, J.C.; Gentile, C.A.; Baer, J. Do gifted student writers and creative writing experts rate creativity the same way? *Gift. Child Q.* **2005**, *49*, 260–265.
82. Baer, J.; Kaufman, J.C.; Riggs, M. Rater-domain interactions in the Consensual Assessment Technique. *Intl. J. Creativity Problem Solving* **2009**, *19*, 87–92.
83. Baer, J.; Kaufman, J.C.; Gentile, C.A. Extension of the consensual assessment technique to nonparallel creative products. *Creativity Res. J.* **2004**, *16*, 113–117.

84. Kaufman, J.C.; Baer, J.; Cropley, D.H.; Reiter-Palmon, R.; Sinnett, S. Furious activity vs. understanding: How much expertise is needed to evaluate creative work? *Psychol. Aesthet. Creativity Arts* **2013**, *7*, 332–340.
85. Kaufman, J.C.; Baer, J.; Cole, J.C. Expertise, domains, and the Consensual Assessment Technique. *J. Creative Behav.* **2009**, *43*, 223–233.
86. Kaufman, J.C.; Baer, J.; Cole, J.C.; Sexton, J.D. A comparison of expert and nonexpert raters using the Consensual Assessment Technique. *Creativity Res. J.* **2008**, *20*, 171–178.
87. Beghetto, R.A.; Kaufman, J.C.; Baer, J. *Teaching for Creativity in the Common Core Classroom*; Teachers College Press: New York, NY, USA, 2014.
88. Dollinger, S.J. Creativity and conservatism. *Personal. Individual Differ.* **2007**, *43*, 1025–1035.
89. Priest, T. Self-evaluation, creativity, and musical achievement. *Psychol. Music* **2006**, *34*, 47–61.
90. Baer, J. Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Res. J.* **1994**, *7*, 35–46.
91. Myford, C.M. The Nature of Expertise in Aesthetic Judgment: Beyond Inter-Judge Agreement. Unpublished Doctoral Dissertation, University of Georgia, Athens, Greece, 1989.
92. Kaufman, J.C.; Evans, M.L.; Baer, J. The American Idol Effect: Are students good judges of their creativity across domains? *Empir. Stud. Arts* **2010**, *28*, 3–17.
93. Horng, J.S.; Lin, L. The development of a scale for evaluating creative culinary products. *Creativity Res. J.* **2009**, *21*, 54–63.
94. Reiter-Palmon, R.; Mumford, M.D.; Threlfall, K.V. Solving everyday problems creatively: The role of problem construction and personality type. *Creativity Res. J.* **1998**, *11*, 187–197.
95. Kaufman, J.C.; Lee, J.; Baer, J.; Lee, S. Captions, consistency, creativity, and the consensual assessment technique: New evidence of validity. *Think. Skills Creativity* **2007**, *2*, 96–106.
96. Kaufman, J.C. Counting the muses: Development of the Kaufman-Domains of Creativity Scale (K-DOCS). *Psychol. Aesthet. Creativity Arts* **2012**, *6*, 298–308.
97. Kaufman, J.C.; Baer, J. Intelligent testing with Torrance. *Creativity Res. J.* **2006**, *18*, 99–102
98. Green, G.P.; Kaufman, J.C. (Eds.). *Videogames and Creativity*; Academic Press: San Diego, CA, USA, 2015.
99. Hamlen, K. Children's choices and strategies in video games. *Comput. Hum. Behav.* **2011**, *27*, 532–539.
100. Jackson, L.A.; Witt, E.A.; Games, A.I.; Fitzgerald, H.E.; von Eye, A.; Zhao, Y. Information technology use and creativity: Findings from the children and technology project. *Comput. Hum. Behav.* **2012**, *28*, 370–376.
101. Kim, Y.J.; Shute, V.J. Opportunities and challenges in assessing and supporting creativity in video games. In *Videogames and Creativity*; Green, G.P., Kaufman, J.C., Eds.; Academic Press: San Diego, CA, USA, 2015; pp. 101–123.
102. Shute, V.J.; Ventura, M. *Measuring and Supporting Learning in Games: Stealth Assessment*; MIT Press: Cambridge, MA, USA, 2013.
103. Shute, V.J.; Ventura, M.; Kim, Y.J. Assessment and learning of informal physics in Newton's Playground. *J. Educ. Res.* **2013**, *106*, 423–430.

104. Dumont, R.; Viezel, K.D.; Kohlhagen, J.; Tabib, S. A review of Q-interactive assessment technology. *NASP Commun.* **2014**, *43*, 8–12.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).