Finding Creative Potential on Intelligence Tests via Divergent Production

James C. Kaufman1, Scott Barry Kaufman2, and Elizabeth O. Lichtenberger3

Abstract
Assessing creative potential using a comprehensive battery of standardized tests requires a focus on how and why an individual responds in addition to how well they respond. Using the “intelligent testing” philosophy of focusing on the person being tested rather than the measure itself helps psychologists form a more complete picture of an examinee, which may include information about his or her creative potential. Although most aspects of creativity are not present in current individually based IQ and achievement tests, one exception is divergent production. Although still poorly represented, some subtests show great potential for tapping into divergent production, and hence provide some insight into creativity. The research on the relationship between measures of intelligence and creativity is discussed in this article. The authors also propose a way to use individually administered cognitive and achievement batteries to extract information about an individual's divergent production and general creative potential.

Résumé
Evaluer le potentiel créatif en utilisant une batterie de tests standardisés implique de s'intéresser, non seulement aux bonnes réponses d'une personne, mais aussi au « comment » et au « pourquoi » de ses réponses. Utiliser une philosophie de « testing intelligent », s'interessant à la personne évaluée plutôt qu'à la mesure elle-même, permet au psychologue de construire un portrait plus complet du sujet en y incluant des informations sur son potentiel créatif. La plupart des aspects de la créativité ne

1California State University at San Bernardino
2New York University
3Carlsbad, CA

Corresponding Author:
James C. Kaufman, Learning Research Institute, California State University at San Bernardino, Department of Psychology, 5500, University Parkway, San Bernardino California, CA 92407
Email: jkaufman@csusb.edu
Clinicians who desire to develop a “complete” picture of a child or adult from an assessment must look beyond the scores on an intelligence test, test of achievement, or other cognitive test. Indeed, referral questions that clinicians try to answer via a comprehensive assessment battery often demand a focus on how an individual responds, or why an individual responds in a certain manner, in addition to how well they respond. In particular, when referral questions include the question of how creative an individual is or whether an individual’s tendency toward creativity influences their performance, then the clinician must focus on how the problems are solved and why individuals respond in the manner that they do. The need to assess a broader scope of abilities (such as creativity) than those measured by traditional tests is also driven by recent neuropsychological research (Delis et al., 2007).

When assessing creativity as part of a comprehensive evaluation, clinicians benefit from following an intelligent testing approach. The primary tenet of the concept of “intelligent testing” is that when an examiner administers an IQ test or other standardized cognitive assessment instruments, the focus should be on the person being tested, not the measure itself (A. S. Kaufman, 1979, 1994, 2009, 2009b; Lichtenberger & Kaufman, 2009). The examiner can help the child or adult being evaluated by observing and interpreting a wide range of test behaviors, making inferences about observed problem-solving strategies, and, ultimately, interpreting the test profile within the context of pertinent background information about the person, clinical behaviors observed during the evaluation, and the latest theories and research in the field of cognitive processing. Just as the test is administered individually, so, too, should the test interpretation be geared to the specific person being evaluated. The entire assessment process is, in effect, an experiment conducted with \( N = 1 \). In other words, in an experiment, the empirical results are of limited value until they are interpreted and discussed in the context of research and theory by a knowledgeable researcher. Similarly, the empirical outcomes of an IQ test are often meaningless until put into context by the examiner (Lichtenberger & Kaufman, 2009).
Although some researchers purport that interpretation of test scores via a cross-validation approach is “contradicted by the research literature” (Watkins & Canivez, 2004, p. 137), it is important to note that the interpretation of IQ data and data related to creativity cannot be conducted in a vacuum. Anastasi and Urbina (1997), and many others have aptly pointed out that sound principles of assessment practice require more than the mechanical application of profile analysis techniques. Lichtenberger and Kaufman (2009, pp. 133-143) and Flanagan and Kaufman (2009) address the concerns of those who advocate against profile interpretation (favoring instead to focus instead on only global IQs; for example, McDermott, Fantuzzo, Glutting, Watkins, & Baggaley, 1992). In the process of extracting information about creativity from performance on IQ tests, it is necessary to look beyond a global IQ (or g). Indeed, important data would be lost if analysis stopped at the global ability level (Flanagan & Kaufman, 2009; Nyden, Billstedt, Hjelmquist, & Gillberg, 2001).

Thus, many aspects of psychology are brought into play to analyze and interpret a cluster of scores in the context of accumulated research, theory, and clinical practice. All this information is added to the store of what was known about the client before the testing sessions even began. The accumulated background information and reasons for referral are all part of what is included in forming conclusions and preparing treatment and remedial suggestions that attempt to answer the referral questions.

What can clinicians do when a referral question requires them to find evidence of a person’s creative potential from a comprehensive assessment? Using the intelligent testing philosophy, clinicians may be able to find evidence of creativity within the administered IQ subtests. To demonstrate how this may be done, this article will first discuss current conceptions of creativity and common assessments used to measure creative potential. Next will be a summary of the research on the relationship between measures of intelligence and creativity and then a final proposition of a way to make use of individually administered cognitive and achievement batteries to extract information about individual’s creative potential, based on the guiding philosophy of “intelligent testing.”

What Is Creativity?

One of the earliest conceptions of creativity was Guilford’s (1950, 1967) structure of intellect model. Guilford placed creativity into a larger framework of intelligence as he attempted to organize all of human cognition along three dimensions. The first dimension was called “operations” and simply meant the mental gymnastics needed for any kind of task. The second dimension, “content,” referred to the general subject area. The third dimension, “product,” represented the actual products that might result from different kinds of thinking in different kinds of subject matters. With 5 operations, 4 contents, and 6 products, Guilford’s model had 120 different possible mental abilities (Guilford, 1967). He later expanded the model to include 180 different abilities, though the 120 abilities model is the one more often studied (Guilford, 1988).
One of the Guilford operations (or thought processes) is divergent production—analyzing responses to questions with no obvious, singular answer (such as “What would happen if people no longer needed sleep?”). Guilford originally described divergent production as consisting of four specific abilities. These involve fluency, flexibility, originality, and elaboration. A common test used to measure divergent production is the Unusual Uses Test (Guilford, Merrifield, & Wilson, 1958), where participants are asked to list all the uses of a familiar object, such as a brick. In this context, fluency is measured by the number of responses, or sheer quantity. Flexibility is measured by the variety of different categories or concepts that are evoked. Elaboration is measured by the level of descriptiveness of each use. Originality is measured by uniqueness of response in comparison to the other participants. Modern researchers use the more broad term “divergent thinking” to describe what Guilford referred to as divergent production.

Divergent thinking is an important cognitive process that is associated with future creative achievement (Runco, 2010). However, creative achievement requires not only the ability to produce divergent ideas but also the ability to discern which ideas are useful or are most appropriate to a particular goal. As such, the term “creativity” has expanded over the years to cover more than just divergent thinking. Most research and theory-based definitions of “creativity” boil down to two components. First, creativity must represent something different, new, or innovative (Baer, 1993; J. C. Kaufman & Sternberg, 2007; Sternberg, Kaufman, & Pretz, 2002). This component is most related to divergent production ability. Second, creativity must also be appropriate to the task at hand. A creative response is useful and relevant. This component may draw forth more on processes relating to reasoning ability or general cognitive ability (Sligh, Conners, & Roskos-Ewoldsen, 2005). Indeed, J. C. Kaufman and Sternberg add that creative ideas and products should also represent an element of high quality.

Furthermore, even though the focus in this article is on individual creative cognition, the individual is only one vantage point in which creativity can be assessed. A complete understanding of creativity is had by including the broader sociocultural context. Plucker, Beghetto, and Dow (2004) expand on the concept of creativity by stating: “Creativity is the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context” (p. 90).

**How Is Creative Potential Typically Measured?**

Most standardized creativity tests tend to measure divergent production, which, as just noted, is only part of creativity. Perhaps the best-known tests are the Torrance Tests of Creative Thinking (TTCT; Torrance, 1974, 2008), with both verbal and figural subtests. The TTCT is based on Guilford’s concept of divergent production. Recently, flexibility was dropped from the most recent version of the figural test because of its high correlation to fluency (Hébert, Cramond, Neumeister, Millar, & Silvian, 2002). Replacing flexibility were two new categories, abstractness of titles, and resistance to premature closure (Torrance, 2008).
Despite its popularity, there are many questions about the TTCT’s predictive validity. Some claim that the procedures for interpreting scores are not supported by factor analysis (Clapham, 1998, 2004; Heausler & Thompson, 1988; Hocevar, 1979a-b); others point to how easy it is to change scores with different instructions (Lissitz & Willhoft, 1985). But the TTCT has continued to be the most popular, especially in deciding who gets into gifted programs (Hunsaker & Callahan, 1995).

Another popular method of measuring creativity is the Consensual Assessment Technique (CAT; see Amabile, 1996; Baer, Kaufman, & Gentile, 2004), in which participants are asked to create something, such as a poem or a collage. These products are then evaluated, without group discussion and with each item compared only to other items, by appropriate experts. Who are appropriate experts? Experts typically can have specialized knowledge of a field (such as a published poet evaluating creative writing), experience into the group being studied (such as an elementary school teacher evaluating children’s work), or insight into the cognitive processes being studied (such as psychologists). Although these ratings are (necessarily) subjective, research has found that experts show strikingly high levels of agreement (Amabile, 1982, 1996; Baer, 1993), even across different groups of experts (Baer et al., 2004). The definition of what constitutes an appropriate expert has been extended to advanced graduate students in psychology (J. C. Kaufman, Lee, Baer, & Lee, 2007) and gifted novices (J. C. Kaufman, Gentile, & Baer, 2005). However, though some studies persist in using novice raters (see J. C. Kaufman & Baer, in press), regular novices are not appropriate experts for most domains (J. C. Kaufman, Baer, & Cole, 2009; J. C. Kaufman, Baer, Cole, & Sexton, 2008).

The CAT is currently used as a research tool but is rarely used in any type of psychoeducational assessment. The benefits are substantial; for one, assessing actual creative products has high face validity. The problems are also substantial; given the findings (J. C. Kaufman et al., 2009, 2008) that typical novices are not good raters for the CAT, the cost and time involved with obtaining appropriate experts and soliciting their judgments is quite high. J. C. Kaufman et al. (2008) suggest that some domains may be better suited than others for widespread use, and future work could try to train novices to approximate expert ratings. Some initial work has already been done on a modified version of the CAT (Dollinger & Shafran, 2005). If teachers are considered to be experts in certain domains and for certain age groups (i.e., fourth-grade writing), then it would be certainly possible to use the CAT methodology on smaller-scale assessments.

How Are IQ and Creative Potential Related?

Although some studies have demonstrated that the strength of intelligence and creative potential in predicting creative achievement differs depending on the domain (S. B. Kaufman, 2009; MacKinnon, 1962; Roe, 1953; Simonton, 2009), researchers tend to agree that at least some degree of both general intelligence and domain-general creative thinking contributes to creative achievement. Most studies that look at creativity and intelligence use divergent thinking tests (such as the TTCT) or other related paper-and-pencil tests also scored for fluency, originality, and so on.
Researchers have generally found that such paper-and-pencil measures are significantly associated with psychometric measures of intelligence. Typically, verbal measures of creativity are more strongly related to intelligence; however, some nonverbal drawing tasks (such as the House–Tree–Person Test) have been shown to work as potential measures of creativity (Eyal & Lindgren, 1977). However, this relationship is not a particularly strong one (see Barron & Harrington, 1981; Kim, 2005). In most of these studies, divergent thinking’s correlation with IQ is maintained up to a certain level of performance on a traditional intelligence test. This traditional research has argued for a “threshold effect,” in which creative potential and psychometric intelligence are positively correlated at low levels of IQ and continue to be positively correlated through IQs of approximately 120. In nearly all of these studies (conducted on both children and adults), in people with higher IQs, the two constructs have been reputed to show little relationship (e.g., Fuchs-Beauchamp, Karnes, & Johnson, 1993; Getzels & Jackson, 1962).

More recently, however, this theory has come under fire. Some studies support the idea that IQ and creativity are substantially related across the entire IQ spectrum, whereas others show negligible relations between IQ and creativity across the IQ spectrum. On the substantially related side of the debate, Preckel, Holling, and Weise (2006) studied the relationship between the Culture Fair Intelligence Test (Cattell & Cattell, 1960) and creativity (as measured through divergent thinking tests) and found high correlations between intelligence and divergent thinking across all levels of intellectual abilities. Along similar lines, Robertson, Smeets, Lubinski, and Benbow (2010) reviewed research showing that SAT scores (both global scores and ability patterns) of adolescents significantly predict their creative achievement decades later.

On the negligible relation side of the debate, Kim (2005), in a meta-analysis of 21 studies, found virtually no support for the threshold theory, showing very small positive correlations (mean correlation of .174) between measures of cognitive ability (designed to measure g) and measures of creativity and divergent thinking. In a later meta-analysis, Kim (2008) found that creative achievement was (unsurprisingly) much more associated with divergent thinking test performance than with global scores on IQ tests.

It is worth noting, however, that nearly all of these studies did not use traditional, individually administered intelligence tests. In Kim’s (2005) meta-analysis, many of the studies were more than 30 years old, and therefore, were conducted using IQ tests that do not reflect current IQ theory. In addition, most of the studies used group tests, either actual IQ tests or SATs, which correlate highly with IQ tests (see Frey & Detterman, 2004). Although group IQ tests serve a strong purpose in research studies, they are not used by most school psychologists in psychoeducational assessment of children who have particular problems (A. S. Kaufman & Lichtenberger, 2006). One result of this tendency is that there is comparatively little research on how creative potential is related to modern intelligence tests that are based on multidimensional theories of intelligence.

One of the few research studies to use an individually administered, full-length modern IQ test was that of Sligh et al. (2005), who used the Kaufman Adolescent and
Adult Intelligence Scale (KAIT; A. S. Kaufman & Kaufman, 1993). The KAIT measures fluid intelligence (Gf) and crystallized intelligence (Gc) as conceptualized by Horn and Cattell (1966). Broadly speaking, Gf is the ability to solve novel problems and Gc is acquired knowledge. Sligh et al. delved deeper into the intelligence–creativity relationship by specifically examining the relationship between Gf and Gc and a measure of actual creative innovation in a sample of college students. More important, their measure of creative innovation, the Finke Creative Invention Task (FCIT; Finke, 1990), separates out the two major components of creativity: creative generation and creative interpretation. First, Sligh et al. found that Gc showed the same moderate and positive relationship to composite creativity as past studies did. In contrast, Gf showed the opposite pattern. Gf and composite creativity were significantly correlated for the high-IQ group, but they were not significantly correlated for people with average IQs.

Most telling, perhaps, is that across the entire IQ range, both crystallized and fluid IQ were correlated more highly with the interpretation component of creativity than the generation component. In fact, in the high-crystallized-IQ group, the correlation between crystallized IQ and the generation component of creativity was .07, and in the average-fluid IQ group, the correlation between fluid IQ and the generation component of creativity was .03. This finding dovetails with prior research using the FCIT, where 99% of the participants were able to generate a usable form in the generation stage of the task, whereas only 33% were able to produce a practical invention/interpretation (Finke, 1990). These findings suggest the importance of teasing apart the two stages of creativity, as each stage may tap into different processes. These findings may also help to resolve discrepancies in the field, in which some studies display a threshold effect, whereas others don’t. In sum, divergent production and interpretation are both important dimensions of creativity, but the processes associated with each dimension may not completely overlap, with the interpretation component relying more on the type of processes measured by traditional measures of IQ, such as the ability to draw inferences and test hypotheses (Finke, Ward, & Smith, 1992).

Such relationships as found in Sligh et al.’s (2005) study, however, may not be consistent across the lifespan. Whereas Gc tends to remain consistent until age 60 or so, Gf peaks in one’s 20s and declines rapidly thereafter (see Lichtenberger & Kaufman, 2009). Creativity has also been shown to decline in later years (Ruth & Birren, 1985). Ideally, this question could be addressed with different samples of varying ages.

An interesting suggestion posed by Batey and Furnham (2006) is that the role of Gf and Gc to creativity may shift across the lifespan of a creative person. Gf, they argue, might be more important in early stages, such as everyday creativity (e.g., J. C. Kaufman & Beghetto, 2009). Conversely, a later-career creator working at more eminent levels of creativity may rely more on Gc, and, one might also add, long-term storage and retrieval (Glr).

Some other studies have given brief IQ tests with measures of creative potential. Vartanian, Martindale, and Matthews (2009), in a broader investigation looking at how people make judgments, administered the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) and a divergent thinking measure (Alternate Uses). They
found no significant correlations between divergent thinking scores and verbal IQ (comparable to \( Gc \); \( r = .08 \), n.s.), performance IQ (comparable to \( Gf \); \( r = -.12 \), n.s.), or full-scale IQ (\( r = -.02 \), n.s.). Cho, Nijenhuis, van Vianen, Kim, and Lee (2010) administered the *Wechsler Adult Intelligence Scale—Revised* (WAIS-R; Wechsler, 1981) and the TTCT. They found that the overall score (\( g \)) was more related to the subcomponent scores of the TTCT (specifically, abstractness of titles, elaboration, and resistance to premature closure in the figural test and flexibility in the verbal test) than to the broader scores of fluency and originality.

\( Gf \) and \( Gc \) have also been studied with group-based intelligence tests, with differing results. Some studies, such as Singh (2006), found no significant relationships between \( Gf, Gc \), and creativity. Batey, Chamorro-Premuzic, and Furnham (2009) found that \( Gc \) was more tied to divergent thinking fluency scores than were \( Gf \), a \( g \)-based ability measure, and personality scores; Cho et al (2010) found comparable results. Similarly, Furnham and Chamorro-Premuzic (2006) found that \( Gc \) was correlated with the personality trait Openness to Experience, which is traditionally also correlated with creativity (Feist, 2010).

Yet Batey, Furnham, and Safiullina (2010) found that \( Gf \)—and not \( Gc \)—predicted a divergent thinking fluency task. Similarly, Batey, Chamorro-Premuzic, and Furnham (2010) found \( Gf \) to be a significant predictor of ideational behavior (which is connected to creativity). Nusbaum and Silvia (2011), following up on Gilhooly, Fioratou, Anthony, and Wynn’s (2007) discussion of divergent thinking as being an executive cognitive function, examined the role that \( Gf \) and strategy use played in divergent thinking tasks. They found that \( Gf \) did predict creativity—but it also predicted who would benefit most from using a new strategy for the divergent thinking tasks. People with high \( Gf \) improved even more with an efficient strategy.

Martindale (1999) proposed a differential relationship between creativity and processing speed, and his theory has been tested out (Dorfman, Martindale, Gassimova, & Vartanian, 2008; Vartanian, Martindale, & Kwiatkowski, 2007). According to Martindale’s model, people who are creative are selective with their speed of information processing. Early in the creative problem-solving stage, they have more defocused attention in which they attend to many different stimuli (including possibly irrelevant concepts). As a result, they are able to process a larger amount of information to be processed (and thereby lowering their speediness). Later, when the problem is better understood, their attention span is shortened (i.e., they filter out the unnecessary information) and their reaction time is quicker. To test this theory, Vartanian, Martindale, and colleagues created two different types of reaction time tasks. One task included potentially distracting information, whereas the other task did not. In both studies, creative potential and processing speed were negatively correlated for interference tasks and positively correlated for noninterference tasks, supporting Martindale’s theory. This theory is reminiscent of Sternberg’s (1981) distinction between global and local planning: Brighter people spend more time in initial global planning, making outlines and coming up with an overall plan of attack so that later they do not have to spend as much time figuring things out on the spot (local
planning). These results are also consistent with the recent dual-process theory of intelligence (S. B. Kaufman, 2009, 2011) that defines intelligence as the ability to switch between explicit and implicit modes of thought depending on the task constraints, as well as the idea of contextual focus (see Gabora, 2003, 2010; Gabora & Kaufman, 2010).

Divergent Production in Tests of Intelligence

Although creativity is much more than divergent production (and divergent thinking), there are no current subtests on contemporary individually administered IQ tests that tap directly into, for example, originality or aesthetic preference or openness to experience. For that reason, the focus of the rest of the article will be restricted to divergent production. Within the literature on theories of intelligence, the strongest conceptual connection to divergent thinking ability can be found in the Cattell–Horn–Carroll (CHC) theory.

CHC theory is a convergence of the Cattell–Horn theory of fluid and crystallized intelligence (Horn & Cattell, 1966; Horn & Hofer, 1992; Horn & Noll, 1997) and Carroll’s three-stratum theory (1993). McGrew (2005) offers a comprehensive treatment of the CHC theory as well as the two theories that were merged to form it. The CHC model proposes 10 different broad factors (Stratum 2) of intelligence:

- **Fluid intelligence** ($G_f$): the ability to apply a variety of mental operations to solve novel problems, ones that don’t benefit from past learning or experience;
- **Quantitative knowledge** ($G_q$): store of acquired knowledge represents the ability to use quantitative information and manipulate numeric symbols;
- **Crystallized intelligence** ($G_c$): the breadth and depth of a person’s accumulated knowledge of a culture and the ability to use that knowledge to solve problems;
- **Reading and writing** ($G_{rw}$): store of knowledge that includes basic reading, reading fluency, and writing skills required for the comprehension of written language and the expression of thought via writing;
- **Short-term memory** ($G_{sm}$): the ability to apprehend and hold information in immediate awareness and then use it within a few seconds;
- **Visual processing** ($G_v$): ability to generate, perceive, analyze, synthesize, store, retrieve, manipulate, transform, and think with visual patterns and stimuli;
- **Auditory processing** ($G_a$): ability to perceive, analyze, and synthesize patterns among auditory stimuli and to discriminate subtle nuances in patterns of sound;
- **Long-term storage and retrieval** ($G_{lr}$): ability to store information in and fluently retrieve new or previously acquired information (e.g., concepts, ideas, items, names) from long-term memory;
• Processing speed (Gs): ability to fluently and automatically perform cognitive tasks, especially when under pressure to maintain focused attention and concentration;
• Decision speed/reaction time (Gt): the immediacy with which an individual can react to stimuli or a task.

Each of these broad factors is comprised of narrow abilities (Stratum 1). In total, 70 such narrow abilities have been identified and classified (Flanagan, McGrew, & Ortiz, 2000; Flanagan, Ortiz, & Alfonso, 2007). Narrow abilities “represent greater specializations of abilities, often in quite specific ways that reflect the effects of experience and learning, or the adoption of particular strategies of performance” (Carroll, 1993, p. 634). Three or more qualitatively different narrow abilities are classified under each broad cognitive ability. For example, the Glr broad ability has 13 qualitatively unique narrow abilities subsumed underneath it, including the following: associative memory (MA), meaningful memory (MM), free recall memory (M6), associational fluency (FA), expressional fluency (FE), ideational fluency (FI), naming facility (NA), word fluency (FW), figural fluency (FF), figural flexibility (FX), sensitivity to problems (SP), originality/creativity (FO), and learning ability (L1).

Creativity was hypothesized to be strongly linked to the CHC broad ability of fluid intelligence (Gf) in the early stages of the Cattell–Horn Gf-Gc theory (Cattell & Butcher, 1968). However, such a relationship is no longer explicitly part of the CHC theory. The current model, based on factor analytic studies by Carroll (1993) and others, includes originality/creativity as a component of long-term storage and retrieval (Glr). According to a recent presentation of CHC (McGrew, 2009), “Some Glr narrow abilities have been prominent in creativity research (e.g., production, ideational fluency, or associative fluency)” (p. 6). In the detailed description of the model, this sentence is the only mention of creativity, originality, or divergent thinking. Fluid intelligence (Gf) is discussed in terms of its relationship to problem solving and coping with novel problems (both considered to be highly related to creativity). Nonetheless, in current discussions of the relationship of creativity to intelligence as presented by the CHC model, the emphasis is on Glr.

Table 1 lists the Stratum 1 (narrow) abilities that are subsumed by the Glr broad ability. What sort of Glr narrow abilities would be relevant for good performance on a divergent production task such as the Unusual Uses Test? Ideational fluency (FI), word fluency (FW), and figural fluency (FF) seem relevant to divergent production by allowing for the open-ended generation of responses. Both word fluency (FW) and figural fluency (FF) are components of ideational fluency (FI), involving either the retrieval of word associations (WF) or figures (FI). Furthermore, word fluency (FW) is very similar to figural fluency (FI) in that both involve the retrieval of word associations, but word fluency (FW) also involves the ability to retrieve associations that conform to certain semantic specifications. The overall ability to generate ideas (either in the form of words or figures) is, therefore, related to one’s performance on the Unusual Uses Test, a common measure of divergent production. The task requires the
Table 1. Narrow Abilities (Stratum 1) Subsumed by the Long-Term Storage and Retrieval (Glr) Broad Ability Factor (Stratum 2) in the Cattell–Horn–Carroll (CHC) Model of Human Cognitive Abilities

<table>
<thead>
<tr>
<th>Ability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associative memory (MA)</td>
<td>Ability to form associations between words that may or may not be meaningfully related to each other</td>
</tr>
<tr>
<td>Meaningful memory (MM)</td>
<td>Ability to recall information that is related to each other in a meaningful way</td>
</tr>
<tr>
<td>Free-recall memory (M6)</td>
<td>Ability to recall, in any order, as many items as possible from a large list of unrelated items that are presented one at a time</td>
</tr>
<tr>
<td>Ideational fluency (FI)</td>
<td>Ability to rapidly produce a series of ideas, words, or phrases related to a specific condition or object. Quantity is emphasized, instead of quality or originality</td>
</tr>
<tr>
<td>Associational fluency (FA)</td>
<td>Ability to produce a series of words or phrases associated in meaning to a word or concept with a limited range of meaning. Quality is emphasized, instead of sheer quantity</td>
</tr>
<tr>
<td>Expressional fluency (FE)</td>
<td>Ability to rephrase an idea without losing its original meaning. Rephrasing is emphasized here, instead of idea generation</td>
</tr>
<tr>
<td>Naming facility (NA)</td>
<td>Ability to produce names for concepts or things when presented with the things or a drawing of it</td>
</tr>
<tr>
<td>Word fluency (FW)</td>
<td>Ability to produce words that have given characteristics</td>
</tr>
<tr>
<td>Figural fluency (FF)</td>
<td>Ability to draw as many things as possible when presented with a set of visual stimulus. Quantity is emphasized here, instead of quality of originality. This is the nonverbal counterpart to ideational fluency</td>
</tr>
<tr>
<td>Figural flexibility (FX)</td>
<td>Ability to change set and deal with a figural problem that requires a variety of approaches to find a solution</td>
</tr>
<tr>
<td>Sensitivity to problems (SP)</td>
<td>Ability to think of a number of different solutions to problems that are practical in nature, such as naming all the uses of a particular tool</td>
</tr>
<tr>
<td>Originality/creativity (FO)</td>
<td>Ability to produce original and unique responses to a given problem and to develop innovative methods for situations where there is no standard convergent way to solve a problem</td>
</tr>
<tr>
<td>Learning abilities (L1)</td>
<td>Ability to learn new material generally. This is the least well-defined ability</td>
</tr>
</tbody>
</table>

ability to restrict associations to only those that solve a particular problem (e.g., “How many uses for a brick can you think of?”).

Word fluency (FW) is also related to associational fluency (FA), sensitivity to problems (SP), and originality/creativity (FO). However, whereas word fluency (FW) is limited to restricting associations to words that have certain semantic properties, associational fluency (FA), sensitivity to problems (SP), and originality/creativity (FO) apply more generally to divergent production. In addition, FA, SP, and FO are the only three Glr abilities that emphasize quality of response, with the latter two specifically emphasizing practicality of response. All of the above-mentioned Glr abilities, though, relate to the fluency and originality aspects of divergent thinking in some way.
Another Stratum 1 cognitive ability that may be relevant to divergent production is writing ability (Grw-WA), which lies beneath the Stratum 2 ability reading/writing (Grw). Grw-WA involves the “ability to write with clarity of thought, organization, and good sentence structure” (Flanagan et al., 2007, p. 283). Although divergent production isn’t explicitly mentioned in this definition, some subtests that purport to measure Grw-WA are open ended and thus provide an opportunity for the assessment of divergent production.

Some Gf and Gc abilities may also be relevant to divergent production. Some Gc items measure a person’s concept formation ability or a person’s ability to evoke a variety of categories for any given word. Therefore, these Gc items may be related to divergent production ability. Although some Gf items may be related to creativity, they may be more related to the usefulness or interpretation aspect of creativity than divergent production, as they tap into an individual’s reasoning abilities (and, specifically, analogical transfer abilities). These reasoning abilities are used in tasks such as problem solving, which is often considered to be a key component of creativity (e.g., Mayer, 2006).

Horn’s last expansion of the Cattell–Horn model (Horn & Blankson, 2005) was constructed after the CHC model was proposed and can be seen as representing his personal conceptualization of the CHC model. Horn and Blankson use the TSR label in lieu of Glr, and five of its six components would seem to be relevant to creativity. DMT (originality), DMC (spontaneous flexibility), Fi (ideational fluency), Fe (expression fluency), and Fa (association fluency) would all seem to be instrumental cognitive processes related to creative thought. As most current IQ tests use CHC theory as a theoretical foundation, however, and not Horn’s expansion, the CHC terminology will be used in this article. Furthermore, since measures of Gc and Gf aren’t typically measured by open-ended means, they do not offer much opportunity for assessing divergent production. Therefore, measures of Gc and Gf will not be further explored in this article.

In sum, it is clear that there are a number of CHC abilities that play a role in creativity—with some Glr and Grw abilities playing more of a role in divergent production. Because a large number of empirical studies support the validity of the CHC theoretical model, many current IQ tests use the theory as a basis (Alfonso, Flanagan, & Radwan, 2005). To date, four major intelligence test batteries explicitly use CHC theory: the Differential Ability Scales-II (Elliott, 2007); the Kaufman Assessment Battery for Children-II (A. S. Kaufman & Kaufman, 2004a); the Stanford-Binet Intelligence Scales—Fifth Edition (Roid, 2003), and the Woodcock-Johnson-III (Woodcock, McGrew, & Mather, 2001). According to Flanagan et al. (2007), “Nearly all comprehensive, individually administered intelligence batteries that are used with some regularity subscribe either explicitly or implicitly to CHC theory” (p. 18).

With the popularity of the CHC theory in modern assessment instruments, one would expect there to be a number of intelligence and academic achievement subtests that measure each of the CHC abilities, including those related to creativity (or specifically divergent production). In reality however, even though most other Stratum 1
abilities are measured well by current IQ tests, very few subtests specifically measure Glr, and if they do measure Glr, they focus specifically on associative memory (MA), ideational fluency (FI), and figural fluency (FF). Furthermore, the tests that do measure Grw-WA aren’t typically scored for the creativity of the response, instead focusing on correct usage of grammar and context.

Table 2 lists the individually administered intelligence and achievement tests that measure at least one crucial component of divergent production\(^2\) and maps the particular subtests onto a Glr or Grw ability that is relevant to divergent production. What is missing? First, it must be noted that the most widely administered test of cognitive ability—the Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003)—has no subtests that measure divergent production. The Stanford-Binet Intelligence Scales—Fifth Edition (Roid, 2003) also does not have any such subtests.

A look at the content of the subtests on other major cognitive ability and achievement tests clearly shows that the originality aspect of divergent production is not assessed by current intelligence tests. In terms of Glr, though tests of associative memory (MA) do exist, these tests simply measure the ability to form associations and are not directly relevant to divergent production. Table 2 does list some tests that measure ideational fluency (FI) and figural fluency (FF). However, these measures that ask for the production of words or figures that are related to another word or figure focus solely on quantity, not rarity. In addition, various measures of Grw-WA do exist, but the open-ended stories are typically not scored for originality.

### Extracting Information About Creative Potential

When a comprehensive evaluation is conducted to answer a set of referral questions that includes queries regarding an individual’s creative potential, there are steps that can be taken to glean as much information as possible from a psychoeducational battery of tests. Following the principles of “intelligent testing” throughout the evaluation, a test administrator should maintain focus on the individual rather than simply examine test scores. During the test selection, administration, and interpretation process, one should use the following: knowledge of intelligence-creativity research, theoretical evidence related to the CHC model and creativity, and clinical skill or experience in working with creative individuals.

**Test selection.** Consider the tasks listed in Table 2 when determining which may be potentially good measures of the client’s creative potential. Bear in mind that some tasks may be able to provide information about other referral questions as well as about divergent production. Creativity can be present in one or more domains (visual vs. verbal, for example); thus, consider that factor when selecting tests to administer.

**Background information/history.** Prior to test administration, when collecting information from the client or the client’s parents and teachers, one should ask questions that will elicit information about the client’s tendency toward creative thinking. For example, does the client come up with multiple solutions to problems at home/school/work? Does the client get answers wrong because they have their own unique way of
Table 2. Divergent Production-Related Subtests and Their Corresponding Component in the CHC Model

<table>
<thead>
<tr>
<th>CHC-Glr Ideational Fluency (Stratum 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test:</strong> Kaufman Test of Educational Achievement—Second Edition (KTEA-II; A. S. Kaufman &amp; Kaufman, 2004b)</td>
</tr>
<tr>
<td><strong>Description:</strong> Measures the ability to name as many words as possible that start with a certain sound or that belong to a certain category</td>
</tr>
<tr>
<td><strong>Woodcock-Johnson-III Tests of Cognitive Ability</strong> (Woodcock, McGrew, &amp; Mather, 2001)</td>
</tr>
<tr>
<td><strong>Test:</strong> Retrieval fluency</td>
</tr>
<tr>
<td><strong>Description:</strong> Measures the ability to name as many things as possible in one minute in each of three specific categories</td>
</tr>
<tr>
<td><strong>NEPSY-II</strong> (Korkman, Kirk, &amp; Kemp, 2007)</td>
</tr>
<tr>
<td><strong>Test:</strong> Verbal fluency (Fl)</td>
</tr>
<tr>
<td><strong>Description:</strong> Measures the ability to generate as many words as possible within a certain time frame, according to specific semantic and phonemic categories</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHC-Glr Figural Fluency (Stratum 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test:</strong> NEPSY-II (Korkman et al., 2007)</td>
</tr>
<tr>
<td><strong>Description:</strong> Measures the ability to generate as many unique designs as quickly as possible by connecting a series of dots</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHC-Glr Writing Ability (WA; Stratum 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test 1:</strong> Woodcock-Johnson-III Tests of Cognitive Ability (Woodcock et al., 2001)</td>
</tr>
<tr>
<td><strong>Description:</strong> Measures the ability to formulate and write a simple sentence by using a set of three words that relates to a given stimulus pictures</td>
</tr>
<tr>
<td><strong>Test 2:</strong> Writing samples</td>
</tr>
<tr>
<td><strong>Description:</strong> Measures the ability to write down sentences in response to different demands</td>
</tr>
<tr>
<td><strong>Test:</strong> Written expression</td>
</tr>
<tr>
<td><strong>Description:</strong> Requires the ability to write letters, words, sentences, and edit text with correct punctuation and capitalization via a storybook format. In addition, the examinee has to provide a writing sample that tells a story based on a verbal or picture prompt</td>
</tr>
<tr>
<td><strong>Diagnostic Achievement Battery-Third Edition</strong> (DAB-3; Newcomer, 2001)</td>
</tr>
<tr>
<td><strong>Test:</strong> Contextual language</td>
</tr>
<tr>
<td><strong>Description:</strong> Requires the ability to write a short story with a beginning, middle, and ending based on three stimulus pictures</td>
</tr>
<tr>
<td><strong>Peabody Individual Achievement Test- Revised/Normative Update</strong> (PIAT-R/NU; Markwardt, 1997)</td>
</tr>
<tr>
<td><strong>Test:</strong> Written expression, Level 2</td>
</tr>
<tr>
<td><strong>Description:</strong> Requires the ability to write a story related to a picture</td>
</tr>
<tr>
<td><strong>Test of Early Written Language—Second Edition</strong> (TEWL-2; Hresko, Herron, &amp; Peak, 1996)</td>
</tr>
<tr>
<td><strong>Test:</strong> Contextual writing</td>
</tr>
<tr>
<td><strong>Description:</strong> Requires the ability to write a story related to a picture</td>
</tr>
<tr>
<td><strong>Test of Children’s Language</strong> (TOCL; Barenbaum &amp; Newcomer, 1996)</td>
</tr>
</tbody>
</table>

(continued)
Table 2. (Continued)

<table>
<thead>
<tr>
<th>Test: Original writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description: Requires the ability to write an original story about animal friends that contains a beginning, middle, and end</td>
</tr>
</tbody>
</table>

*Test of Written Language—Third Edition* (TOWL-4; Hammill & Larsen, 2009)

<table>
<thead>
<tr>
<th>Test 1: Contextual language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description: Requires the ability to write a story paying attention to sentence structure, grammar, and vocabulary</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test 2: Story construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description: Requires the ability to write a story paying attention to the use of prose, action, sequencing, and theme</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test 3: Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description: Requires the ability to write a sentence using a given word</td>
</tr>
</tbody>
</table>

*Wechsler Individual Achievement Test—Second Edition* (WIAT-II; Wechsler, 2001)

<table>
<thead>
<tr>
<th>Test: Written expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description: Requires the ability to write the alphabet, generate and write a list that represent a given category, combine sentences into one, and generate sentences from provided cues</td>
</tr>
</tbody>
</table>

understanding questions that were asked? Does the client struggle on multiple choice tests because he or she deems none of the answers to be good enough (or he or she comes up with other alternative responses)? Does the client have trouble focusing at times because he or she is deep in thought or daydreaming?

**Behavioral observations.** Be an astute observer of an individual’s problem-solving styles and strategies during the assessment. Gather information about how and why an individual responds in a particular way to items. These observations will help you properly interpret responses to tasks that may be related to their creative potential. For example, does an individual try out many different strategies when solving the same type of problem? Does the individual use a trial-and-error problem solving approach? Does the individual use an organized, systematic problem-solving approach? Does the individual tend to elaborate on responses that don’t necessarily demand elaboration? Does the individual get frustrated when a task has a forced-choice response that doesn’t allow for their own original response? Is the individual’s response style slow and methodical or quick (and perhaps impulsive)? Does the individual indicate that he or she prefers structured or less structured tasks?

**Test interpretation.** Although the tests listed in Table 2 are related to the construct of divergent production (including the Glr narrow abilities of ideational fluency, figural fluency, and writing ability), high scores on these tests do not necessarily equate to a high level of creativity. Each of these task’s scores must be interpreted in the context of the examinee’s referral question, background, observed behaviors, and other test results. When evaluating a person’s creative potential using these tasks from standardized cognitive and achievement tasks, it is important to remember that creative
achievement requires not only the ability to produce divergent ideas but also the ability to discern which of the ideas are appropriate to a particular goal.

Is it possible for an examiner to derive an originality score from the data collected from the subtests mentioned that have potential for measuring divergent production? It is possible, though it is important to reinforce the concept that more research is needed before any direct clinical applications would be recommended. One way of getting originality scores is through the CAT mentioned earlier. As mentioned, psychologists are perfectly qualified to serve as experts (Baer et al., 2004; Baer, Kaufman, & Riggs, 2009). Although there would ideally be several experts judging an individual’s work, an educational assessment is not necessarily designed to be a research study. If an initial suggestion about an individual’s creativity is all that is needed to be included in an intervention plan, the psychoeducational diagnostician’s judgment about the originality of a response may be all that is needed. The measures of Grw-WA listed in Table 2 are particularly well suited toward the CAT technique since the responses are open ended and allow the test taker to express his or her unique imagination.

Future Directions for Measuring Creative Potential in IQ Tests

As sketched out in this article, there are a number of ways that divergent thinking ability can be assessed by using standardized IQ tests. The endeavor is still limited, however, by the fact that the Glr abilities most relevant to divergent thinking, such as ideational fluency (FA), sensitivity to problem (SP), and originality/creativity (FO) are not particularly well measured by modern IQ tests, nor are they well defined. Also, though there is enormous potential to measure originality of response in current measures of Grw-WA, that potential has not yet been realized. Test developers can fill the gaps in what current versions of their instruments measure by creatively integrating new tasks that demand original, quality, and unique responses in addition to quantity.

A change to scoring methodology is another means by which test developers to improve measurement of creativity in existing tasks. For example, a method of obtaining scores of originality would be to simply count the number of responses on an ideational fluency (FI) or figural fluency (FF) subtest. This way examiners could judge the remoteness of the associations by comparing an individual’s associative responses to that of other test takers. The statistical rarity of a particular associative response could then be used as a measure of originality, which would provide a complement to the usual scoring method involving sheer quantity. Currently, norms do not exist regarding responses on these tasks, but we believe that such studies should be done so that test administrators can glean originality information much better from the test taker’s responses. When IQ tests are normalized and standardized, researchers should collect data on the responses to the divergent production-related items coming from each examinee. If this information were published, then test examiners could calculate
an originality score in the same way that originality is calculated in other tests of divergent thinking.

The idea that the ability to form remote associations is associated with creativity is grounded in theory. Mednick (1962, 1968) constructed the Remote Associates Test (RAT) that presented triads of words to participants and then asked them to derive the one word related to all three of the previous words. For example, given a triad of Falling, Actor, and Dust, a correct answer would be Star (a falling star; an actor is a star; stardust). This test is based on the premise that creativity involves the ability to make rather remote associations among separate ideas. It is frequently used as a measure of creativity in empirical studies (e.g., Ansburg & Hill, 2003; Isen, Labroo, & Durlach, 2004; White & Shah, 2006) based on the premise that the further apart conceptually two ideas are, the more creative a person must be to see the connection behind them. Therefore, either employing the consensual assessment technique or comparing responses to divergent thinking-related IQ test items to published norms may allow for an assessment of remote association ability. It is worth pointing out, however, that the RAT typically correlates at a highly significant level with verbal IQ (J. C. Kaufman, 2009a). Some critics question whether the RAT measures mental processes greatly different than those measured by IQ tests.

In addition, even though this article has not discussed subtests relating to elaboration (since our focus was on divergent thinking), very few Glr test items allow for an assessment of the practicality of the response. If test developers are interested in incorporating more creativity-relevant abilities into IQ tests (that go beyond divergent production alone), one basic way would be to expand the different types of association-based items. In addition to asking questions that probe the number of associations an individual may create, additional questions could be added that require the participant to make their ideas both original and practical, and scoring instructions can allow the examiner to give points for both. Then, in addition to assessing the number of responses, test administrators can also assess the originality and practicality of the divergent thinking. There are already some subtests, such as the riddles and double-meanings subtests on the KTEA-II (A. S. Kaufman & Kaufman, 2004a) that already require mental associations. Future research may want to look into investigating how these subtests along with the subtests mentioned in the current paper may relate to actual lifetime creative achievement.

In sum, it should be clear that there is enormous untapped potential to assess the creative potential of examinees using the already-existing individually administered standardized tests of general cognitive ability that are grounded in CHC theory. In addition, test developers can consider expanding their tests’ ability to measure creativity by expanding scoring systems and adding tasks that require divergent thinking. By applying the “Intelligent Testing” philosophy to draw a complete picture of the individual being assessed, hopefully it is possible to improve our identification of students who are not only able to produce answers in a convergent fashion but also excel in divergent production and ultimately have potential for the highest levels of creative achievement.
Acknowledgments

The authors would like to thank Mark Batey, Alan S. Kaufman, and Nadeen L. Kaufman for extensive comments and suggestions and Claudine Wierzbicki for help with translation. The authors would like to dedicate this article to the memory of the late Dr. John L. Horn.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Nearly always conceptualized as “g” or full-scale IQ.
2. It should be noted that the flexibility component is no longer tested on the latest version of the Torrance Tests of Creativity (Torrance, 2008) because of its high correlation to fluency (Hébert, Cramond, Spears-Neumeister, Millar, & Silvian, 2002).

References


**Bios**

**James C. Kaufman**, PhD, is an associate professor of psychology at the California State University at San Bernardino, where he directs the Learning Research Institute. He received his PhD from Yale University in cognitive psychology in 2001. His research broadly focuses on nurturing and encouraging creativity. He is the author or editor of 20 books and more than 175 papers. He is the president-elect of American Psychological Association’s Division 10 (Aesthetics, Creativity, and the Arts) and is a founding coeditor of the division’s official journal, *Psychology of Aesthetics, Creativity, and the Arts*. He is also the founding editor of APA’s newest journal, *Psychology of Popular Media Culture*. He received the 2003 Daniel E. Berlyne Award from APA’s Division 10, the 2008 E. Paul Torrance Award from the National Association of Gifted Children, and the 2009 Early Career Research Award from the Western Psychological Association.

**Scott Barry Kaufman**, PhD, is an adjunct assistant professor of psychology at New York University. He researches and writes about the development of talent, intelligence, creativity, imagination, and personality. In addition to publishing scholarly articles, book chapters, and edited books, including *The Psychology of Creative Writing* (with James C. Kaufman) and *The Cambridge Handbook of Intelligence* (with Robert J. Sternberg), he is also associate editor of *The International Journal of Creativity and Problem Solving* and the Sex, Art, and Pop Culture editor of *The Evolutionary Review*. He also blogs for a blog for *Psychology Today* entitled “Beautiful Minds” and blogs for *The Huffington Post*. He is the recipient of the 2008 Frank X. Barron award from Division 10 of the American Psychological Association for his research on the psychology of aesthetics, creativity, and the arts.

**Elizabeth O. Lichtenberger**, PhD, is a licensed clinical psychologist in California, and an author whose works have focused on psychological assessment. In addition to her current professional roles, she has worked as an adjunct faculty member at Alliant International University in San Diego and a researcher at the Laboratory for Cognitive Neuroscience at The Salk Institute for Biological Studies in La Jolla, CA.