

Metaphorically speaking: Cognitive abilities and the production of figurative speech.

By: Roger E. Beaty & Paul J. Silvia

[Roger E. Beaty](#) & [Paul J. Silvia](#) (2013). Metaphorically speaking: Cognitive abilities and the production of figurative speech. *Memory and Cognition*. 41(2). 255-267. [DOI: 10.3758/s13421-012-0258-5](https://doi.org/10.3758/s13421-012-0258-5)

Made available courtesy of Springer Verlag. The definitive version is available at <http://link.springer.com/>.

*****Reprinted with permission. No further reproduction is authorized without written permission from Springer Verlag. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. *****

Abstract:

Figurative language is one of the most common expressions of creative behavior in everyday life. However, the cognitive mechanisms behind figures of speech such as metaphors remain largely unexplained. Recent evidence suggests that fluid and executive abilities are important to the generation of conventional and creative metaphors. The present study investigated whether several factors of the Cattell–Horn–Carroll model of intelligence contribute to generating these different types of metaphors. Specifically, the roles of fluid intelligence (Gf), crystallized intelligence (Gc), and broad retrieval ability (Gr) were explored. Participants completed a series of intelligence tests and were asked to produce conventional and creative metaphors. Structural equation modeling was used to assess the contribution of the different factors of intelligence to metaphor production. For creative metaphor, there were large effects of Gf ($\beta = .45$) and Gr ($\beta = .52$); for conventional metaphor, there was a moderate effect of Gc ($\beta = .30$). Creative and conventional metaphors thus appear to be anchored in different patterns of abilities: Creative metaphors rely more on executive processes, whereas conventional metaphors primarily draw from acquired vocabulary knowledge.

Keywords: metaphor production | intelligence | Carrol model | psychology | psychonomics | cognitive abilities | language | figurative language

Article:

Figurative language is perhaps the most common expression of creativity in everyday life (Carter, 2004). People often use figures of speech like metaphors to describe a vast array of emotions and experiences. Although figurative language pervades human dialogue, our understanding of how people come up with these types of expressions is quite limited. Psycholinguistic research has produced a wealth of knowledge on metaphor comprehension (e.g., Gibbs, 1994; Glucksberg, 2001; Kintsch, 2000; Lakoff & Johnson, 1980), but we are only now

starting to understand how the mind creates metaphors (Chiappe & Chiappe, 2007; De Barros, Primi, Miguel, Almeida, & Oliveira, 2010; Pierce & Chiappe, 2009; Silvia & Beaty, 2012).

Nevertheless, all metaphors are not created equal; they vary in terms of novelty and familiarity. In the present study, we were interested in examining the cognitive processes involved in two types of figurative language. Conventional metaphors are straightforward, often clichéd or idiomatic expressions. Such metaphors are typically highly apt: Their structure represents a comprehensible and appropriate comparison between a topic and a characteristic exemplar (Tourangeau & Sternberg, 1982). For example, the conventional description “life is a journey” entails a one-to-one comparison, one that is familiar and easy to comprehend. Creative metaphors, in contrast, are distinctly original uses of language: They are unique in that neither the creator nor the audience has encountered the metaphor before. Creative metaphors are frequently used in conversation to describe an emotional experience (Carter, 2004), developed to express imagery in literature (Plotnik, 2007), and employed as symbolic tools in several artistic traditions (Kennedy, 2008).

The present research examined the contribution of several cognitive abilities from the Cattell–Horn–Carroll (CHC) model of intelligence (McGrew, 2009) to the generation of conventional and creative metaphors. Modern CHC theory, an emerging consensus on the structure of cognitive abilities based on decades of factor-analytic research (see Carroll, 1993), conceptualizes intelligence as a hierarchical structure of abilities, from narrow (specific abilities, such as inductive reasoning) to broad (global abilities, such as fluid intelligence) to general (g). Recent evidence demonstrates the strong contribution of fluid intelligence—a broad cognitive ability—to the generation of creative metaphors (Silvia & Beaty, 2012). It remains unclear, however, how other aspects of intelligence influence metaphor production. This study thus explores how three broad CHC abilities—fluid intelligence (Gf), crystallized intelligence (Gc), and broad retrieval ability (Gr)—influence how people produce both conventional and creative metaphors.

The property attribution model

Although there are several definitions of metaphor, one prominent description categorizes it as a higher-order term that includes other structures, like similes and analogies (Barnden, 2010; Grady, 2007). Metaphors function as a descriptive mechanism of communication; they describe a specific aspect of a given topic by relating it to a conceptually similar exemplar. Exemplars are often referred to as vehicles, and they embody some level of abstract relation to a referent concept (i.e., the topic). The study of metaphor comprehension has been of interest to cognitive linguists for several decades, and a large body of research has been dedicated to understanding metaphoric structure and function (e.g., Gibbs, 1994; Glucksberg, McGlone, & Manfredi, 1997; Lakoff & Johnson, 1980). Nevertheless, an empirical understanding of how the mind produces figurative language remains elusive.

The property attribution model of metaphor comprehension provides a useful framework for conceptualizing metaphor production. According to this model, composing a metaphor involves making an abstract link between a topic and a vehicle by relating similar characteristics (Glucksberg, 2001). Shared conceptual knowledge between the topic and vehicle must be identified for a metaphor to be comprehensible (Glucksberg et al., 1997). While people search semantic memory for an appropriate vehicle, a superordinate attributive category maintains some characteristics of the topic that can be used to relate to the vehicle. For example, if one were to consider a metaphor for “music,” an attributive category—“something that is healing”—guides the search process en route to an appropriate vehicle (“medicine”).

Several aspects of Glucksberg’s (2001) property attribution model can be adopted to conceptualize the cognitive mechanics of metaphor generation. First, the formation and maintenance of a higher-order attributive category is analogous to Carroll’s (1993) concept of Gr. According to Carroll and others, Gr represents the capacity to fluently extract knowledge from long-term memory (Cattell, 1978; Horn, 1988; McGrew, 2005). Tasks developed to assess retrieval ability typically require people to generate members from a given category on the basis of a presented cue (e.g., “list synonyms for the word good”). Considered in the context of Glucksberg’s model, one can see an apparent parallel between attributive categories and broad retrieval ability: Searching memory for a candidate vehicle to attribute to a specific topic seems much like the selective retrieval processes associated with Gr.

Furthermore, exercising top-down oversight of the metaphor generation process has been shown to recruit executive abilities associated with Gf (Silvia & Beaty, 2012). Previous research demonstrates Gf’s considerable association with working memory capacity (Kane et al., 2004) and implicates this ability in other controlled processes, such as directing attention during complex cognitive tasks (Heitz, Unsworth, & Engle, 2005) and managing interference from task-irrelevant information (Unsworth, 2010). Central to the attribution model described by Glucksberg (2001) is the process of relating two otherwise semantically unrelated concepts (e.g., lawyers and sharks). One must prevent the literal or adjectival information closely linked to the topic and vehicle from interfering with the goal of making a figurative connection (e.g., some lawyers can be predatory, but they do not share the physical characteristics of sharks). We would thus expect fluid and executive abilities to facilitate the search process by maintaining the task goal in mind and inhibiting inapt associates that compete for activation in memory (Gernsbacher, Keysar, Robertson, & Werner, 2001).

Conventional metaphor production

An interest in metaphor generation has reemerged in the past decade, with several researchers attempting to identify the underlying cognitive processes involved (Chiappe & Chiappe, 2007; De Barros et al., 2010; Pierce & Chiappe, 2009; Silvia & Beaty, 2012). In a series of experiments, Chiappe and Chiappe administered measures of executive function and a series of metaphor tasks. In Experiment 1, participants completed a working memory task (listening

span), a measure of inhibitory control (Stroop task), and a metaphor comprehension task that they designed (see Chiappe & Chiappe, 2007). Participants were split into high and low working memory span on the basis of performance on the listening span task. High-spans produced better metaphor interpretations—scored for quality on a 3-point scale by two raters—and did so at a faster rate than low-spans. Intrusion errors on the Stroop task were negatively correlated with the quality of metaphor interpretations and the length of time it took participants to generate these interpretations.

The second experiment assessed metaphor generation with a fill-in-the-blank completion task. For the metaphor task, participants had 15 min to complete 24 figurative statements, and they were given property descriptions to relate to each vehicle (e.g., “Some jobs are _____”; Property: something that is confining and constraining and can make you feel like you’re just putting in time). Two raters scored responses on a 6-point scale for aptness. Several executive tasks were administered, including measures of working memory (listening span), verbal fluency (generating first names, foods and drinks, and animals), and vocabulary knowledge (Peabody Picture Vocabulary Test; PPVT). Performance on the PPVT explained most of the variance in metaphor quality ($R^2 = .17$); listening span scores explained a smaller yet significant portion of variance ($R^2 = .10$). Commonality analysis was used to determine distinct contributions of variance from these independent measures. The unique proportion of variance contributed by listening span reduced to 2.9 %, with the remaining variance attributed to a shared contribution along with vocabulary knowledge.

Similar to Experiment 2, the third experiment assessed the role of working memory and vocabulary knowledge in the composition of conventional metaphors. Working memory was measured with listening span, digit span forward, and digit span reverse tasks. Participants completed the PPVT, verbal fluency tasks, the metaphor generation task, and the Magazine Recognition Questionnaire—a measure of familiarity with printed media—to assess one component of general knowledge. As a set, commonality analysis revealed that PPVT and listening span scores explained 31.3 % in metaphor quality, of which a majority (26.6 %) was contributed by the PPVT. A second commonality analysis including print exposure and listening span tasks showed a similar pattern ($R^2 = .29$), with listening span explaining 9.1 % of unique variance. Digit span tasks were not included in the reported analysis, since performance on these tasks was weakly correlated with metaphor quality. Taken together, the authors interpreted the results from these three experiments as an indication that crystallized knowledge and executive abilities each influence the process of metaphor production.

Creative metaphor production

Investigations of conventional metaphor demonstrate how people construct simple figurative statements that are straightforward and easily interpreted. They have been limited to analyzing singular, discrete vehicles that are produced in response to fill-in-the-blank tests (e.g., Chiappe & Chiappe, 2007; Christensen & Guilford, 1963; De Barros et al., 2010; Taylor, 1947). In some

cases, metaphor completion tasks have included additional constraints by essentially providing a definition of the to-be-produced vehicle (e.g., Chiappe & Chiappe, 2007; Pierce & Chiappe, 2009). These studies contribute to a greater understanding of conventional thinking, but they have several limitations for studying creative cognition.

Recent evidence suggests that fluid intelligence is essential to the creative thought process (Beaty & Silvia, in press; Benedek, Franz, Heene, & Neubauer, 2012; Nusbaum & Silvia, 2011; Silvia & Beaty, 2012; Vartanian, Martindale, & Kwiatkowski, 2003). Several executive mechanisms have been shown to facilitate individual differences in creative thinking, such as controlling attention during idea generation (Vartanian, 2009; Zabelina & Robinson, 2010), implementing effective cognitive search strategies (Gilhooly, Fioratou, Anthony, & Wynn, 2007; Nusbaum & Silvia, 2011), and switching between semantic categories in memory (Nusbaum & Silvia, 2011). Considering the substantial contribution of fluid and executive abilities to domain-general creative thinking, one might expect these mechanisms to support similar types of cognition.

Silvia and Beaty (2012) examined the contribution of fluid intelligence to creative metaphor quality. Participants were presented with two different prompts and were asked to describe past emotional experiences using a metaphor. The first prompt asked people to “think of the most boring high-school or college class that you’ve ever had. What was it like to sit through?” For the next prompt, participants were asked to “think about the most disgusting thing you ever ate or drank. What was it like to eat or drink it?” Responses were scored by three raters on a 5-point scale using subjective scoring (Silvia et al., 2008). Six measures of inductive reasoning—primarily nonverbal and visual-spatial—were administered to assess fluid intelligence. Participants also completed the Five Factor Inventory, which measures the Big Five factors of personality (McCrae & Costa, 1997). Structural equation models revealed a large effect of fluid intelligence in predicting the creative quality of metaphors (standardized $\beta = .49$), and this effect remained large when personality was added to the model. Together, personality and fluid intelligence explained 35 % of the variance in creative metaphor quality.

The present research

In the present study, we explored the contribution of cognitive abilities to the generation of creative and conventional metaphors. Our previous study (Silvia & Beaty, 2012) demonstrated that fluid intelligence strongly predicts the creative quality of metaphors. One aim of the present research was to extend this finding. Studies of conventional metaphor suggest that executive abilities such as working memory contribute to the generation of conventional metaphors (Chiappe & Chiappe, 2007; Pierce & Chiappe, 2009). Are executive abilities equally as important for generating both creative and conventional metaphors? Since working memory and fluid intelligence are closely related constructs (Kane et al., 2004; Süß, Oberauer, Wittman, Wilhelm, & Schulze, 2002), we examined the contribution of fluid intelligence to both creative and conventional metaphors.

Another primary goal of this project was to explore how different cognitive abilities contribute to conventional and creative metaphor. Specifically, we were interested in testing aspects of the CHC model of intelligence (Carroll, 1993; McGrew, 2005). Past research has shown that Gf is broadly important to divergent thinking (Nusbaum & Silvia, 2011) and creative metaphor (Silvia & Beaty, 2012). Considering the results from the conventional metaphor literature (e.g., Chiappe & Chiappe, 2007; De Barros et al., 2010), one would expect Gf to be important for generating apt metaphors as well. In the present research, we examined the contribution of Gf to figurative language, with the goal of determining its relative importance to both types of metaphor.

We were also interested in examining the degree to which general knowledge influences metaphor production. Figurative statements involve vocabulary and other knowledge about the world, and people certainly vary in terms of acquired information (Kan, Kievit, Dolan, & van der Mass, 2011). Carroll (1993) referred to this acquired knowledge as crystallized intelligence (Gc), a higher-order factor that “develops through the investment of general intelligence into learning through education and experience” (p. 599). To what extent does acquired knowledge contribute to creative and conventional thinking?

Furthermore, past research indicates that Gr—the capacity to fluently recall concepts from long-term memory (Cattell, 1978)—supports the creative thought process (Gilhooly et al., 2007; Silvia, Beaty, & Nusbaum, 2012). Taken in the context of the present study, the ability to retrieve knowledge from memory in an efficient and fluent manner should play an important role as well. More specifically, the generation of retrieval cues seems to be a key aspect of the selective search process. The notion that coming up with a metaphor engages selective retrieval mechanisms fits particularly well with the property attribution model of Glucksberg et al. (1997). Creating and deploying a superordinate attributive category—one that functions to guide the semantic search for appropriate descriptive vehicles—greatly resembles Carroll’s (1993) conceptualization of Gr.

In this study, participants completed several assessments that measure three factors of the CHC model of intelligence: fluid intelligence, crystallized intelligence, and broad retrieval ability. We administered the conventional metaphor task of Chiappe and Chiappe (2007) as well as the creative metaphor task used in our prior study. Considering the property attribution model of Glucksberg (2001) as a framework for metaphor generation, several executive processes should be important. For example, maintaining an attributive category in mind while searching semantic memory and managing interference from inapt, obvious, and adjectival information should require executive resources captured by Gf.

Finally, we were interested in the contribution of personality variables to metaphor production. Our previous study, using the NEO Five Factor Inventory (FFI; Costa & McCrae, 1992), found a large effect of openness to experience on creative metaphor (Silvia & Beaty, 2012). Openness consistently correlates with creativity (Feist, 1998; McCrae, 1987; Silvia, Nusbaum, Berg, Martin, & O’Connor, 2009) and with intelligence (Ashton, Lee, & Vernon, 2000; Goff &

Ackerman, 1992), so openness is a potential “third variable.” The effects of personality traits on metaphor production are interesting in their own right, particularly to researchers in the large field of personality and creativity, but they also afford a test of the incremental validity of the CHC abilities.

Method

Participants

The sample consisted of 222 undergraduate students from the University of North Carolina at Greensboro (156 women, 66 men). Participation was voluntary, and students received credit toward a research option in a psychology class for their involvement in the study. Students who indicated that English was not their primary language were excluded from analysis ($n = 18$). Multivariate outlier tests revealed one highly outlying case, which was then excluded. In addition, data from participants who exhibited disengagement with the study were withheld from the analysis ($n = 12$). Exclusion criteria included finishing the hour-long study in less than 30 min, “clicking through” tasks (as evidenced by experimenter observation notes and data analysis), frequently text-messaging, and holding conversation with other participants. The final sample consisted of 191 students (135 women, 56 men). The self-identified ethnic composition of the final sample was 57 % European American, 29 % African American, 5 % Hispanic/Latino, 4 % Native American, 3 % Asian American, and 3 % undeclared.

Procedure

The study was carried out in a group setting, with the number of participants ranging from 1 to 8 per session. Students filled out consent forms and were briefed on the study procedure by an experimenter. Following informed consent, students completed metaphor tasks, several cognitive tasks, and some personality questionnaires. MediaLab v2010 software, run on standard Windows-based desktop PCs, was used to administer all measures in the study.

Metaphor tasks

Conventional metaphor generation task

The metaphor task from Chiappe and Chiappe (2007) was used to assess individual differences in conventional metaphor generation. Twenty-four metaphor prompts were selected from the list of items. Task instructions—including metaphor examples and explanations—were taken verbatim from the Chiappe and Chiappe test manuals and were presented to participants in the present study. Several figurative statements provided examples during the instructions phase, along with accompanying explanations of their structure. Each item presented a topic and property description, followed by a fill-in-the-blank metaphor stem (e.g., “Come up with a metaphor that conveys that some jobs are confining and constraining, and make you feel like you are just putting in time”; “Some jobs are _____”).¹

Participants were asked to complete each statement with a vehicle that appropriately related to the topic (e.g., “jails”). If they were unable to think of a vehicle, they were instructed to type “I don’t know” into the response dialogue box. Students had 15 min to complete the 24 metaphors. Following the procedure of Chiappe and Chiappe (2007), two raters scored the vehicles for aptness, using a 6-point scale (from 0 to 5). Highly apt responses successfully attributed the vehicle to the topic vis-à-vis the specified property and received a score of 5. Instances where participants could not think of an appropriate vehicle received a score of 0. The remainder of the scale (i.e., 1–4) was applied to vehicular responses that related to the topic with varying degrees of aptness and abstractness. Raters were blind to each other’s scores and to the participants’ other data. For each rater, the 24 items were averaged to get a continuous score.

Creative metaphor generation task

Following the conventional metaphor task, participants were asked to describe two past experiences with a metaphor. The aim of this task was to assess creative thinking and participants’ ability to come up with a unique response. Instructions included definitions and examples of different types of metaphors (e.g., simile, metaphor, and compound metaphor). The experimenter informed students that they could work on the task for as long as they liked. Following the instructions phase, participants read the first of two metaphor prompts: “Think of the most boring high-school or college class that you’ve ever had. What was it like to sit through?” Examples of metaphoric stems were provided as potential starting points (e.g., “Being in that class was like . . .”). The second prompt stated: “Think about the most disgusting thing you ever ate or drank. What was it like to eat or drink?” Stems were also provided for this prompt (e.g., “Eating that ____ was like . . .”).

Instructions for this task included several descriptive terms to distinguish the characteristics of a creative response. We asked participants to “be creative” and “to come up with something that is clever, humorous, original, compelling, or interesting.” Previous studies of divergent thinking have demonstrated that instructions to “be creative” typically result in more unique responses (Christensen, Guilford, & Wilson, 1957; Harrington, 1975; Niu & Liu, 2009). For the present study, it was particularly important to discriminate between conventionality and creativity, especially since participants had just completed the conventional metaphor task. The examples of creative metaphors included figurative statements that were more elaborate and interesting than the conventional metaphor samples.

Responses were scored by three raters using subjective scoring (Amabile, 1982; Christensen et al., 1957; Silvia, 2011; Silvia et al., 2008). Each metaphor received a score of 1 (not at all creative) to 5 (very creative). Raters were trained to score responses on the basis of three criteria: remoteness, novelty, and cleverness. Remoteness reflects the conceptual distance of the metaphor—the extent to which the vehicle related to the topic abstractly. Novelty reflects the originality of the response: clichés and common idioms received a low score (e.g., “It was like watching paint dry”). Cleverness reflects the degree to which the response was funny, witty, or

interesting. Although there were several criteria, each response received a single score from each rater. The subjective scoring method was used in our prior study of metaphor (Silvia & Beaty, 2012) and has been shown to be a reliable assessment of divergent thinking (Silvia, 2011). Similar to conventional metaphor scoring, raters were unaware of each other's scores and of the participants' other data.

Fluid intelligence (Gf)

Letter sets task

This task presents a series of five letter sets with four letters in each set. Four of the sets follow a specific rule, such as vowel–consonant–vowel–consonant (e.g., ACIF). One of the letter sets does not follow the rule, and the goal is to identify this set. Participants must choose the correct answer from a list of five answer choices. The task included 16 items and was timed for 4 min (Ekstrom, French, Harman, & Dermen, 1976).

Cattell Culture Fair Intelligence task

The series completion task was adopted from Cattell's Culture Fair Intelligence Test (Cattell & Cattell, 1961/2008). Each task item has a row of boxes. Patterns within the boxes changed according to a specific rule, and the objective was to determine the successive element from a list of answer choices. Participants had 3 min to complete 13 problems.

Paper folding task

This task assesses visual-spatial reasoning ability, which covaries strongly with fluid intelligence (Kane et al., 2004). Each item presented a square piece of paper followed by a series of images that represented the paper being folded and punched with holes. Participants were to imagine the paper being unfolded and determine the final state of the paper from a series of answer choices. The task included ten items and was timed for 3 min (Ekstrom et al., 1976).

Broad retrieval ability (Gr)

Three verbal fluency tasks were administered, and participants had 1 min to enter as many responses as they could think of for each. The first prompt required students to generate synonyms for the word good, an assessment of associational fluency (Carroll, 1993). After the synonyms task, two other verbal fluency tasks were administered: a word fluency task (words that start with the letter M) followed by an ideational fluency task (occupations). The letter M task required participants to generate different words that start with M; the occupations prompt required the generation of types of jobs (Carroll, 1993).

Responses for all of the fluency tasks received two scores: overall output and adjusted output. The total number of responses—regardless of accuracy and repetition—was summed to calculate

the overall output score. Adjusted output, which removed repetitions and invalid responses, was used in the analysis.

Crystallized intelligence (Gc)

Vocabulary

Two tests from the ETS Kit of Factor-Referenced Cognitive Tasks assessed vocabulary knowledge: the Advanced Vocabulary Test II (18 items, 4 min) and the Extended Range Vocabulary Test (24 items, 4 min; Ekstrom et al., 1976). Questions from both tests presented a target word with four to five answer choices. Participants were asked to choose the word that best described the target word.

General knowledge tests

Multiple choice questions were administered from three general knowledge domains: general biology, literature, and American history. While there are many areas of specialized knowledge, the intent of the general knowledge tests was to obtain a measure of knowledge in common fields of study. Since the sample was made up of undergraduate students, it was reasonable to assume that coursework in these fields was completed by nearly all participants prior to college. Literature questions were compiled from various subject texts. The history and biology test items came from high school advanced placement (AP) study guides. Questions were taken from practice tests and chosen on the basis of a criterion of medium difficulty. Participants had 10 min to answer 30 questions (10 from each domain).

Personality

Following the cognitive tasks, participants completed the NEO Five Factor Inventory (FFI; Costa & McCrae, 1992). The questionnaire consists of 60 items and measures five factors of personality: openness to experience, neuroticism, extraversion, agreeableness, and conscientiousness (McCrae & Costa, 1997). Items consist of statements that reflect one of the five factors, and participants rate the statements on a 5-point scale (1 = strongly disagree; 5 = strongly agree). We measured personality to assess each factor's unique contribution to metaphor production and to determine the incremental validity of the intelligence variables.

Results

Model specification

The data were analyzed with structural equation models, using Mplus 6.12 with maximum likelihood estimation. Dependent variables included conventional and creative metaphors, and independent variables were the cognitive abilities. Prior to analyzing the full structural model, we ran confirmatory factor analyses (CFAs) to test the loadings of all measures and model fit. The three factors of intelligence were modeled as latent variables—Gr, Gc, and Gf—indicated by

scores on their respective tasks (see Fig. 1). Creative and conventional metaphors were also modeled as latent variables, indicated by the scores of the raters. Table 1 lists the descriptive statistics and correlations; Table 2 displays the correlations between the latent variables.

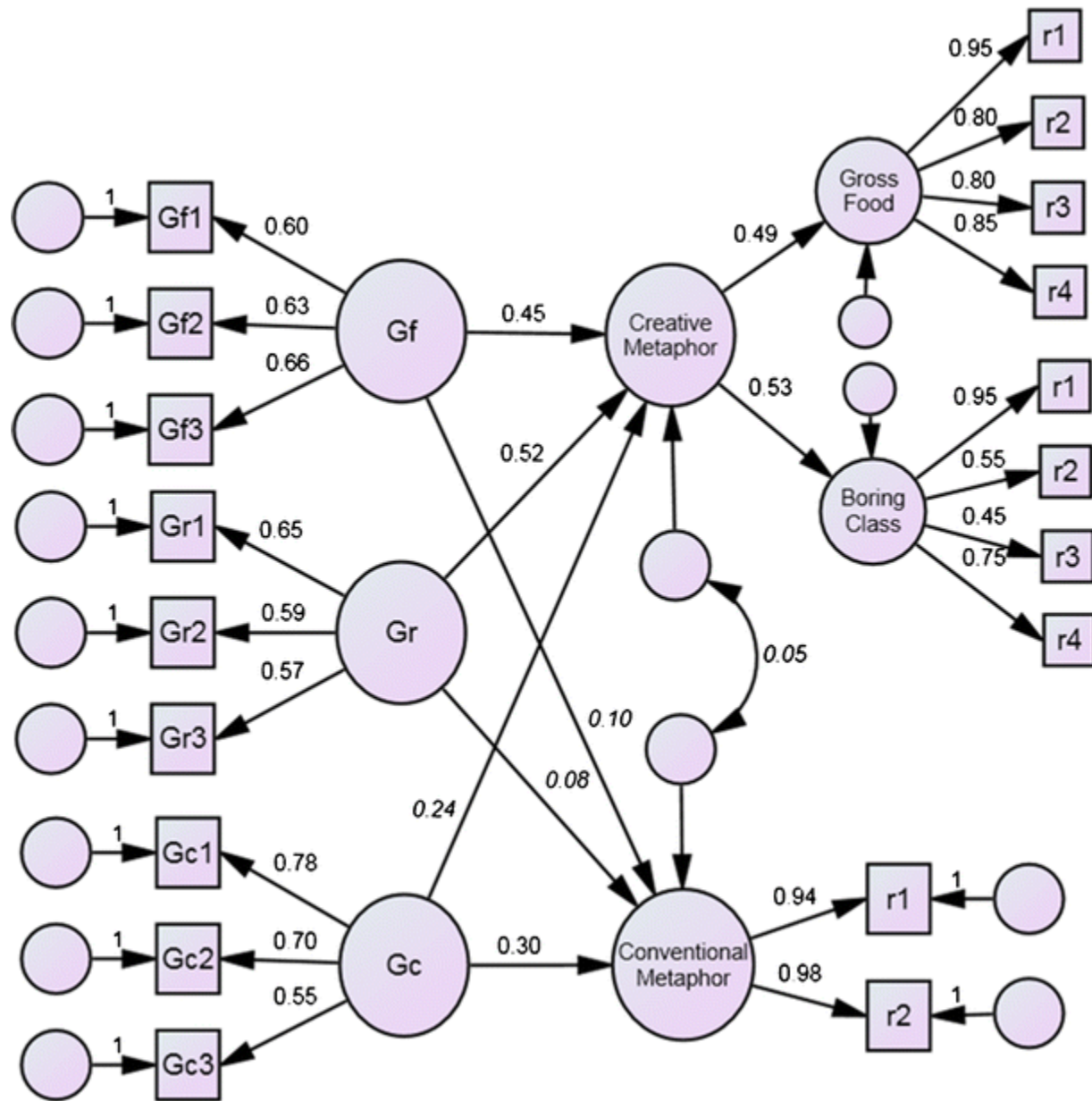


Fig. 1

Effects of cognitive abilities on creative and conventional metaphors. The creative metaphor rating indicators are ordinal and, thus, do not have residual variances. Italicized values are not significant at $p < .05$. Abbreviated indicator labels for the intelligence factors represent the following tasks: Gf1 = Cattell Series Completion, Gf2 = Paper Folding, Gf3 = Letter Sets; Gr1 = Jobs, Gr2 = Letter *M* Words, Gr3 = “Good” Synonyms; Gc1 = Extended Vocabulary, Gc2 =

Advanced Vocabulary, and Gc3 = U.S. History. The correlations between Gf, Gr, and Gc were omitted for clarity. The correlation between Gr and Gf was .27, between Gc and Gf was .48, and between Gc and Gr was .38

Table 1

Correlations and descriptive statistics

	<i>M</i>	<i>SD</i>	Min, Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Gross metaphor: Rater 1	1.13	0.38	1, 3	1																		
2. Gross metaphor: Rater 2	1.29	0.67	1, 4	.43	1																	
3. Gross metaphor: Rater 3	1.37	0.65	1, 4	.48	.63	1																
4. Gross metaphor: Rater 4	1.25	0.54	1, 3	.53	.58	.63	1															
4. Boring metaphor: Rater 1	1.18	0.46	1, 4	.21	.16	.13	.13	1														
5. Boring metaphor: Rater 2	1.26	0.56	1, 4	.00	.02	-.01	-.04	.09	1													
6. Boring metaphor: Rater 3	1.33	0.62	1, 4	.05	.05	.14	.15	.35	.25	1												
8. Boring metaphor: Rater 4	1.33	0.55	1, 3	.00	.08	.13	.08	.23	.34	.52	1											
9. Conventional metaphor avg.: Rater 1	3.80	0.64	.96, 4.96	.00	.12	.05	.00	.06	.05	.14	.05	1										
10. Conventional metaphor avg.: Rater 2	3.51	0.61	.88, 4.83	.06	.15	.09	.05	.83	.05	.14	.03	.92	1									
11. Gf: Series completion	7.87	1.66	1, 11	.01	.09	.21	.05	-.03	-.07	.22	.02	.22	.20	1								
12. Gf: Paper folding	5.76	2.16	0, 10	.06	-.02	.22	.07	-.04	-.07	.34	.07	.15	.15	.36	1							
13. Gf: Letter sets	8.31	2.62	2, 14	.04	-.02	.26	.01	-.07	-.04	.28	.14	.11	.12	.40	.40	1						
14. Gr: Jobs	12.20	3.06	3, 23	.20	.04	.14	.15	.11	.23	.08	.12	.17	.15	.09	.06	.14	1					
15. Gr: LetterM	14.61	3.61	3, 24	.12	.08	.11	.07	.10	.04	.16	.19	.11	.15	.06	.07	.12	.38	1				
16. Gr: Synonyms for <i>good</i>	7.92	2.98	2, 20	.19	.13	.27	.23	.20	.00	.15	.13	.07	.07	-.04	.11	.21	.35	.33	1			
17. Gc: Extended vocabulary	10.08	3.37	1, 19	.17	.08	.16	.15	.13	.07	.37	.21	.23	.23	.20	.31	.30	.24	.09	.19	1		
18. Gc: Advanced vocabulary	7.79	2.50	1, 14	.20	.12	.13	.12	-.01	-.09	.11	.06	.24	.26	.17	.23	.10	.18	.07	.09	.54	1	
19. Gc: History	4.35	1.96	0, 10	.14	.08	.01	.02	.02	.00	.10	.15	.28	.31	.16	.03	.17	.23	.11	.05	.38	.43	1

Table 2

Correlations between the latent variables

	1	2	3	4	5
1. Creative metaphor	1				
2. Conventional metaphor	.41	1			
3. Fluid intelligence	.82	.31	1		
4. Broad retrieval ability	.82	.28	.40	1	
5. Crystallized intelligence	.79	.43	.60	.50	1

Intelligence

Our first model analyzed the factor structure of the three intelligence variables. The variances of Gf, Gc, and Gr were fixed to 1. A CFA of the specified model suggested good fit: $\chi^2(41) = 58.47, p = .047$; CFI = .94; SRMR = .049; RMSEA = .047 (90 % CI: .012, .073). However, the Gc variable showed mixed loadings for its five indicators. Specifically, the biology and literature measures loaded poorly on the Gc factor. Further analysis of internal consistency revealed low alphas for biology (Cronbach's $\alpha = .10$) and literature ($\alpha = -.07$). The history test loaded moderately on the Gc factor and showed modest but adequate internal consistency ($\alpha = .50$). The two vocabulary tests were the most robust indicators of our Gc factor, consistent with past research linking vocabulary knowledge with crystallized intelligence (Carroll, 1993). Factor loadings for the Gf and Gc latent variables were moderate in magnitude. The revised CFA (see Fig. 1), with the biology and literature variables excluded from the model, showed good fit: $\chi^2(24) = 41.08, p = .016$; CFI = .94; SRMR = .049; RMSEA = .061 (90 % CI: .026, .092). As a comparison, we considered the fit of an implausible model: A general intelligence variable was specified with all nine intelligence tests serving as indicators. As was expected, this model showed considerably worse fit: $\chi^2(27) = 134.128, p < .001$; CFI = .633; SRMR = .101; RMSEA = .144 (90 % CI: .120, .168).

Creative metaphor

Scores from the two creative metaphor tasks were specified as categorical variables. Similar to our past research analyzing subjective ratings of verbal creativity tasks (e.g., Nusbaum & Silvia, 2011), we found the distribution of scores to be highly skewed. A majority of responses received low scores from the four raters, and very few were coded at the upper end of the scale (i.e., 4s and 5s). This level of skew violates the assumption of multivariate normality and leads to issues with model convergence (Kline, 2011). Modeling the scores as ordinal handles the skewed ratings by estimating the likelihood of each score (i.e., 1–5) without assuming multivariate normality (Nusbaum & Silvia, 2011; Silvia & Beaty, 2012).

Each metaphor—"gross food" and "boring class"—was specified as a lower-order latent variable indicated by the four raters' scores (see Fig. 1). We also specified a higher-order creative

metaphor factor with the “gross food” and “boring class” variables as indicators. The paths were constrained to be equal so that the higher-order variable could be identified. The variance of this higher-order factor was fixed to 1. One notable issue with analyzing categorical variables in structural equation models is the shortage of fit statistics suitable to assess model fit. However, the structural model reached convergence, which is an indication of an admissible solution (Kline, 2011). We have also specified similar structural equation models in our previous study of cognitive ability and creative metaphor (Silvia & Beaty, 2012).

We estimated the reliability of the subjective ratings using generalizability theory (Cronbach, Gleser, Nanda, & Rajaratnam, 1972), an extension of classical reliability statistics that is ideal for crossed and nested designs (Brennan, 2001; Shavelson & Webb, 1991). In our study, for example, we have two tasks crossed by four raters. Using EduG 6.1 (Cardinet, Johnson, & Pini, 2010), we specified a design in which the tasks were fixed and the participants and raters were random. The G coefficient—a holistic estimate of reliability with the same scale and interpretation as Cronbach’s alpha—was .78, indicating a good level of score reliability.

Conventional metaphor

Conventional metaphor was modeled as a latent variable indicated by the two raters’ scores. The lower-order rating variables were constrained to be equal, and the higher-order factor’s variance was fixed to 1 for model identification. Similar to the procedure described in Chiappe and Chiappe (2007), we computed an average of each participant’s total score on all 24 items. The interrater reliability for the two raters was quite high (Cronbach’s $\alpha = .97$).

Intelligence and creative metaphor

We analyzed the direct effects of Gf, Gc, and Gr on the quality of creative metaphors. Figure 1 depicts the structural model and standardized effects; Table 3 lists details for the regression effects. As was expected, fluid intelligence strongly predicted creative metaphors ($\beta = .45, p = .017$). This effect size can be interpreted as “large” using the benchmarks of .10 for small, .30 for medium, and .50 for large (Cohen, 1988). Interestingly, the magnitude of this effect was similar to that in our previous study ($\beta = .49$; Silvia & Beaty, 2012). Gr had a similarly large effect on metaphor creativity ($\beta = .52, p < .001$), and Gc had a moderate but nonsignificant effect ($\beta = .24, p = .206$).

Table 3

Summary of the regression effects

Model	Creative metaphor				Conventional metaphor			
	Standardized beta	P value	95 % confidence interval	R ²	Standardized beta	P value	95 % confidence interval	R ²
1. CHC								
Gf	.45	.017	.08, .83	.20	.10	.371	-.11, .30	.01
Gr	.52	.001	.23, .82	.27	.08	.469	-.13, .29	.00

Gc	.24	.206	-.13, .62	.05	.30	.005	.09, .51	.09
2. Personality								
Neuroticism	.14	.366	-.16, .44	.02	.00	.933	-.16, .15	.00
Extraversion	-.26	.120	-.59, .07	.06	-.08	.313	-.23, .08	.00
Openness to experience	.61	.000	.38, .85	.37	.14	.051	.00, .29	.02
Agreeableness	.22	.188	-.11, .55	.05	.03	.730	-.13, .18	.05
Conscientiousness	-.08	.614	-.41, .24	.00	-.16	.045	-.32, .00	.03
3. CHC and personality								
Gf	.45	.013	.10, .80	.20	.09	.405	-.12, .30	.01
Gr	.55	.000	.25, .85	.30	.05	.677	-.18, .28	.00
Gc	-.02	.945	-.47, .44	.00	.29	.035	.02, .56	.08
Neuroticism	.04	.743	-.19, .27	.00	.00	.996	-.15, .15	.00
Extraversion	-.32	.056	-.64, .01	.10	-.02	.872	-.18, .22	.00
Openness to Experience	.35	.006	.10, .60	.12	.03	.752	-.13, .19	.00
Agreeableness	.14	.286	-.12, .41	.02	-.02	.796	-.18, .14	.00
Conscientiousness	.12	.359	-.14, .37	.01	-.09	.274	-.25, .07	.01

$n = 191$

Intelligence and conventional metaphor

For the conventional metaphor task, the structural model found a significant effect of Gc on metaphor ratings ($\beta = .30, p = .005$). However, the direct effects of Gf ($\beta = .10, p = .371$) and Gr ($\beta = .08, p = .469$) were small and nonsignificant (see Fig. 1 and Table 3). These findings point to a minor influence of executive abilities but a larger influence of general knowledge in conventional metaphor production. It is worth noting that we estimated the effects of intelligence on conventional metaphor using structural equation modeling, whereas Chiappe and Chiappe (2007) used ANOVA models. While the present study differs in this regard, analyzing latent variables typically yields larger effects than do observed variables (Kline, 2011; Silvia, 2008).²

Personality and metaphor production

We assessed the role of personality in producing both types of metaphor. First, we entered the five factors of personality (Costa & McCrae, 1992) into a model as predictors of creative metaphor. The factors were specified as observed variables to simplify the large model. Table 3 displays a summary of the standardized regression coefficients. Consistent with past research on personality and creativity (Batey & Furnham, 2006; Feist, 1998; McCrae, 1987), openness to experience had a substantial effect on creative metaphor quality ($\beta = .61, p < .001$). All other personality variables had small and nonsignificant effects. A second model included the personality variables as predictors of creative metaphor, along with the three intelligence factors. This allowed for a test of incremental validity of the CHC variables. Openness again predicted creative metaphors ($\beta = .35, p = .006$). The effect sizes for Gf ($\beta = .45$) and Gr ($\beta = .55$) remained stable, while the effect of Gc on creative metaphors was reduced to zero ($\beta = -.02, p = .945$). Previous studies have shown moderate correlations between Gc and openness (Ashton et al., 2000; Goff & Ackerman, 1992), so the diminished effect is likely due to their shared variance.

Regarding conventional metaphor, a model specified with the personality variables as predictors yielded small effects for conscientiousness ($\beta = -.16, p = .045$) and openness ($\beta = .14, p = .051$). Including personality in a model with Gf, Gr, and Gc predicting conventional metaphor quality yielded similar effects for the IQ variables (see Table 3). However, the effects of conscientiousness ($\beta = -.09, p = .274$) and openness ($\beta = .03, p = .752$) were decreased. Taken together, the inclusion of personality factors in a model with the intelligence variables influenced the weak coefficients for conventional metaphor most significantly, whereas the large effects of Gf and Gr on creative metaphors were largely unchanged.³

Discussion

The present study provides several insights into the nature of metaphor production. Our results demonstrate the differential contribution of cognitive abilities to the generation of creative and conventional metaphors. One goal of the present research was to replicate the results from our previous study of creative metaphor (Silvia & Beaty, 2012). Specifically, we again tested the hypothesis that fluid intelligence would predict the creative quality of metaphors. Structural equation models revealed this effect to be almost identical to that in our prior analysis. Including additional factors of intelligence in the present analysis allowed us to extend our previous study and take a closer look at other underlying mechanisms involved in creative ideation.

Recent studies have reported close links between conventional metaphor generation and higher-order executive processes (e.g., Chiappe & Chiappe, 2007; Pierce & Chiappe, 2009). In our study, however, fluid intelligence had a small effect on conventional metaphor production. This observation does not entirely contradict the results from Chiappe and Chiappe's work; they measured the effect of executive mechanisms with assessments of working memory and controlled attention, although the effects on conventional metaphor quality were small and nonsignificant in some cases. Nevertheless, considering the strong association between fluid intelligence and working memory capacity (Kane, Hambrick, & Conway, 2005), one would expect at least a modest relation between Gf and conventional metaphor. Our study bolstered the likelihood of observing this relationship by analyzing latent variables, which remove measurement error and typically yield larger effect sizes (Kline, 2011).

Our results suggest a less important role of executive processes in conventional metaphor generation, since Gf and Gr showed small effects in our structural equation model. But it is worth noting that we did not explicitly manipulate the conventionality of topics and vehicles, as did Pierce and Chiappe (2009) in their follow-up study of metaphor production. They found that the conventionality of topic–category pairs predicted vehicle aptness: Participants tended to converge on an apt solution when the attributive category was closely linked with the topic. Such a manipulation was not central to this study; our main goal was to contrast the different cognitive abilities that influence two distinct types of metaphors—conventional and creative—and our analysis suggests that these recruit different processes. We nonetheless encourage future

researchers to explore the nuances and underlying mechanics of these distinct forms of figurative language.

Vocabulary knowledge accounted for the most variance in Chiappe and Chiappe's (2007) study, and Gc had a medium effect on conventional metaphor production in our study. Thus, producing conventional metaphors might primarily draw upon prior knowledge and minimally recruit executive resources. On the other hand, the results for crystallized intelligence and creative metaphor quality are not as clear, considering that our latent Gc factor did not significantly predict the creative quality of metaphors and had essentially no effect after controlling for personality.

Another notable result from the present analysis involved the strong relation between Gr and creative metaphor. Crafting a novel metaphor should recruit selective retrieval processes, so we expected Gr to play a role, although the magnitude of this effect was greater than we anticipated ($\beta = .52$). In light of past research, however, one might expect Gr to contribute to creative ideation. For example, Carroll's (1993) CHC model of intelligence includes divergent thinking as a facet of Gr. Creative metaphor production could be considered a close cousin of divergent thinking, although the tasks used to measure these constructs differ in their demands (e.g., the elaboration of a single response in creative metaphor generation vs. the generation of several alternate uses for an object in divergent thinking). Nevertheless, the theoretical basis of broad retrieval ability should be developed further to better understand its function in metaphor production.

The function of retrieval ability might not be surprising when considered in terms of Glucksberg's (2001) property attribution model. Glucksberg posited that the process by which we comprehend a figurative statement includes the creation and maintenance of a superordinate attributive category that serves to relate the topic of a metaphor to a vehicle (Glucksberg et al., 1997). Such a mechanism has been adopted in recent models of metaphor generation (e.g., Chiappe & Chiappe, 2007; Silvia & Beaty, 2012). In a typical selective retrieval task, one must deploy a given search cue (e.g., "words that start with the letter M") to extract relevant information from long-term memory. This type of targeted search process also fits with our conceptualization of metaphor use. Producing a novel metaphor—one that meets certain abstract criteria—should recruit selective retrieval mechanisms vis-à-vis a higher-order attributive category. In our experiment, participants had to form an attributive category for "things that are gross" and selectively retrieve exemplars that satisfied these criteria. The close resemblance in task demands thus helps to explain the large effect of Gr on creative metaphor quality.

Perhaps the most important avenue for future research to pursue is to dig beneath the global effects of intelligence to uncover the fundamental cognitive processes involved in metaphor production. Much is known about the nature of fluid intelligence, such as its strong association with working memory and inhibitory control (Kane et al., 2004), but how does this ability specifically influence the process of generating figurative language? Many abilities fall under the

umbrella of executive process (Friedman & Miyake, 2004), and they may have different influences on metaphor production, as they have been shown to have with respect to divergent thinking (Benedek et al., 2012). Using the property attribution model as a guiding framework, future work should continue to examine the ways in which executive processes facilitate creative metaphor production.

Conclusion

Figurative language is a common mode of communication, but the cognitive processes that support both creative and conventional metaphor use are not well understood. The present analysis revealed that higher-order mechanisms associated with executive processes predicted the quality of creative metaphors, while crystallized knowledge predicted peoples' ability to generate conventional metaphors. Taken together, this study provides new evidence for the differential contribution of intellectual abilities to metaphor production, and it extends the study of creative cognition within the CHC intelligence framework.

References

- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997–1013.
- Ashton, M., Lee, K., Vernon, P., & Jang, K. (2000). Fluid intelligence, crystallized intelligence, and the openness/intellect factor. *Journal of Research in Personality*, 34, 198–207.
- Barnden, J. A. (2010). Metaphor and metonymy: Making their connections more slippery. *Cognitive Linguistics*, 21, 1–34.
- Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, 132, 355–429.
- Beaty, R. E., & Silvia, P. J. (in press). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts*.
- Benedek, M., Franz, F., Heene, M., & Neubauer, A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences*, 53, 480–485.
- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer.
- Cardinet, J., Johnson, S., & Pini, G. (2010). *Applying generalizability theory using EduG*. New York: Routledge.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

- Carter, R. (2004). *Language and creativity: The art of common talk*. New York: Routledge.
- Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum.
- Cattell, R. B., & Cattell, A. K. S. (1961/2008). *Measuring intelligence with the Culture Fair Tests*. Oxford, UK: Hogrefe.
- Chiappe, D. L., & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of Memory and Language*, 56, 172–188.
- Christensen, P. R., & Guilford, J. P. (1963). An experimental study of verbal fluency factors. *British Journal of Statistical Psychology*, 16, 1–26.
- Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology*, 53, 82–88.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah: Lawrence Erlbaum Associates.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa: Psychological Assessment Resources.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- De Barros, D. P., Primi, R., Miguel, F. K., Almeida, L. S., & Oliveira, E. P. (2010). Metaphor creation: A measure of creativity or intelligence? *European Journal of Education and Psychology*, 3, 103–115.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2, 290–309.
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control processes: A latent variable analysis. *Journal of Experimental Psychology: General*, 133, 101–135.
- Gernsbacher, M. A., Keysar, B., Robertson, R. R., & Werner, N. K. (2001). The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language*, 45, 433–450.

Gibbs, R. W., Jr. (1994). *The poetics of mind: Figurative thought, language, and understanding*. New York: Cambridge University Press.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology*, 98, 611–625.

Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms*. New York: Oxford University Press.

Glucksberg, S., McGlone, M., & Manfredi, D. (1997). Property attribution in metaphor comprehension. *Journal of Memory and Language*, 36, 50–67.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessing typical intellectual engagement. *Journal of Educational Psychology*, 84, 537–552.

Grady, J. E. (2007). Metaphor. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 188–213). New York: Oxford University Press.

Harrington, D. M. (1975). Effects of explicit instructions to “be creative” on the psychological meaning of divergent thinking test scores. *Journal of Personality*, 43, 434–454.

Heitz, R. P., Unsworth, N., & Engle, R. W. (2005). Working memory capacity, attention control, and fluid intelligence. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 61–77). New York: Sage.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 645–685). New York: Plenum.

Kan, K., Kievit, R. A., Dolan, C., & van der Mass, H. (2011). On the interpretation of the CHC factor Gc. *Intelligence*, 39, 292–302.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 66–71.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133, 189–217.

Kennedy, J. M. (2008). Metaphors in art. In R. W. Gibbs (Ed.), *Cambridge handbook of metaphor and thought* (pp. 447–461). Cambridge: Cambridge University Press.

- Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, 7, 257–266.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, 4, 100–107.
- McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509–516.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 136–181). New York: Guilford.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1–10.
- Niu, W., & Liu, D. (2009). Enhancing creativity: A comparison between effects of an indicative instruction “to be creative” and a more elaborate heuristic instruction on Chinese student creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 3, 93–98.
- Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence*, 39, 36–45.
- Pierce, R. S., & Chiappe, D. L. (2009). The roles of aptness, conventionality, and working memory in the production of metaphors and similes. *Metaphor and Symbol*, 24, 1–19.
- Plotnik, A. (2007). *Spunk and bite: A writer’s guide to bold, contemporary style*. New York: Random House Reference.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park: Sage.
- Silvia, P. J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, 44, 1012–1021.
- Silvia, P. J. (2011). Subjective scoring of divergent thinking: Examining the reliability of unusual uses, instances, and consequences tasks. *Thinking Skills and Creativity*, 6, 24–30.
- Silvia, P. J., & Beaty, R. E. (2012). Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence*, 40, 343–351.

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2012). Verbal fluency and creativity: Higher-order and lower-order contributions of broad retrieval ability (Gr) to divergent thinking. Manuscript under review.

Silvia, P. J., Nusbaum, E. C., Berg, C., Martin, C., & O'Connor, A. (2009). Openness to experience, plasticity, and creativity: Exploring lower-order, higher-order, and interactive effects. *Journal of Research in Personality*, 43, 1087–1090.

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, 2, 68–85.

Süß, H. M., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence*, 30, 261–288.

Taylor, C. W. (1947). A factorial study of fluency in writing. *Psychometrika*, 12, 239–262.

Tourangeau, R., & Sternberg, R. J. (1982). Understanding and appreciating metaphors. *Cognition*, 11, 203–244.

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence*, 38, 255–267.

Vartanian, O. (2009). Variable attention facilitates creative problem solving. *Psychology of Aesthetics, Creativity, and the Arts*, 3, 57–59.

Vartanian, O., Martindale, C., & Kwiatkowski, J. (2003). Creativity and inductive reasoning: The relationship between divergent thinking and performance on Wason's 2-4-6 task. *Quarterly Journal of Experimental Psychology*, 56, 641–655.

Zabelina, D. L., & Robinson, M. D. (2010). Creativity as flexible cognitive control. *Psychology of Aesthetics, Creativity, and the Arts*, 4, 136–143.

Footnotes

1 We considered this task a measure of *conventional metaphor production*, defined as the ability to generate a vehicle term that aptly fits the constraints of an attributive category. This term is used for the sake of distinguishing Chiappe and Chiappe's (2007) task from our creative metaphor task. For the conventional task, responses should be familiar: Most participants should converge on the same small range of responses. For the creative task, responses should be unfamiliar: They needn't fit the bounds of an attributive category, and most participants should give unique responses. All metaphors should be more or less apt, of course. But for our purposes, conventional metaphors were assessed for aptness, and creative metaphors were assessed for novelty.

2 We did not measure typing speed in this study, but in hindsight, it seemed like it could be potentially confounded with performance on the speeded Gr tasks. In a subsequent study of 16 Gr tasks ($n = 131$; Silvia et al., 2012), we found near-zero correlations between typing speed and the measures of broad retrieval ability used in the present study (synonyms for *good*, $r = .058$; occupations, $r = .013$; words that start with the letter *M*, $r = -.022$).

3 Out of curiosity, we explored alternate scoring methods to see whether they would influence the model results. One such method recoded responses ranging from zero to three as 0, and those from four to five as 1, effectively scoring vehicles as “apt” or “inapt.” This procedure yielded a similar pattern of effect sizes in regard to the intelligence variables (e.g., Gf, $\beta = .12$, $p = .277$; Gc, $\beta = .34$, $p = .002$), with the most notable change being a decrease in the effect of Gr ($\beta = -.04$, $p = .724$).