Making creative metaphors: The importance of fluid intelligence for creative thought.

By: Paul J. Silvia & Roger E. Beaty

Paul J. Silvia & Roger E. Beaty (2012). Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence, 40(4),* 343-351.

Abstract:

The relationship between intelligence and creativity remains controversial. The present research explored this issue by studying the role of fluid intelligence (Gf) in the generation of creative metaphors. Participants (n = 132 young adults) completed six nonverbal tests of Gf (primarily tests of inductive reasoning) and were then asked to create metaphors that described a past emotional experience. The metaphors were rated for creative quality. Latent variable models found that Gf explained approximately 24% of the variance in metaphor quality (standardized beta = .49), consistent with the view that creative ideation engages executive processes and abilities. The effect of Gf remained substantial after including personality (the Big Five factors) in the model. The discussion considers implications for the debate over intelligence and creativity as well as for the cognitive abilities involved in metaphor production.

Keywords: creativity | metaphor | psychology | intelligence | creative thought | fluid reasoning | personality | openness to experience | Cattell-Horn-Carroll model

Article:

1. Making creative metaphors: the role of fluid intelligence in creative thought

Are intelligent people more creative, or are intelligence and creativity independent abilities? This question is one of the enduring controversies in the psychology of creativity (Kaufman, 2009 and Wallach and Kogan, 1965). In the present work, we take a new slant on this problem by examining the role of fluid intelligence (Gf) in the production of creative metaphors. This work extends studies of creative cognition to a new domain, provides further support for our view that intelligence is central to creative thought (Nusbaum & Silvia, 2011a), and contributes to the emerging literature on how people make metaphors (Chiappe & Chiappe, 2007).

2. The creativity-and-intelligence controversy

In the psychology of creativity, most reviews of the creativity-and-intelligence controversy have concluded that creativity and intelligence are distinct abilities with minor overlap (e.g., Batey and Furnham, 2006, Kaufman and Plucker, 2011, Kim et al., 2010 and Runco, 2007). Since Wallach and Kogan's (1965) landmark work on this topic, research has typically found that creative cognition—usually measured with divergent thinking tasks—covaries modestly with intelligence. A recent meta-analysis of the relationship between intelligence and divergent thinking found an overall effect of r = .17 (Kim, 2005).

At the same time, many contemporary researchers have found that there are good reasons to expect stronger relationships between intelligence and creative cognition. Generating creative ideas—ideas that are both novel and appropriate to the purpose at hand—requires identifying and implementing strategies for idea generation (Gilhooly et al., 2007 and Nusbaum and Silvia, 2011a), exerting control over attention and thought (Vartanian, 2009, Zabelina and Robinson, 2010 and Zabelina et al., 2012), making decisions and refining initial ideas (Finke et al., 1992, Gabora, 2005 and Vartanian, 2011), and inhibiting obvious and inapt ideas (Nusbaum & Silvia, 2011a).

If this view of creative cognition is right, then fluid and executive abilities should be central to the creative process. But past reviews and Kim's (2005) meta-analysis conclude otherwise, so an executive interpretation of creative thought is understandably controversial. We have suggested that some common methods in creativity research have obscured and deflated the true relationship between intelligence and creative cognition (Nusbaum and Silvia, 2011a, Silvia, 2008a and Silvia, 2008b). First, analyzing latent variables instead of observed variables yields higher effects (Silvia, 2008a). Second, analyzing higher-order abilities—such as fluid intelligence (Gf) or g—yields stronger relationships than analyzing lower-order abilities and individual tasks (Silvia, 2008a). Third, and perhaps most important, newer methods of creativity assessment will yield larger effects. Our past work (Silvia et al., 2009a and Silvia et al., 2008) has contended that the usual ways of assessing divergent thinking have serious problems. Divergent thinking tasks can be scored in many ways (see Plucker, Qian, & Wang, 2011), but the most common ways are to score the number of responses (fluency; e.g., Batey et al., 2009, Preckel et al., 2006 and Preckel et al., 2011) or to score the number of responses given by no one else (uniqueness or originality; Wallach & Kogan, 1965). Uniqueness is confounded with fluency (Silvia, 2008b), and it has an unusual sample dependency—it shrinks as the sample size rises (Silvia, 2011 and Silvia et al., 2008)—that makes it poorly suited for large-sample research.

As an alternative, we have suggested subjective ratings of creativity, which have been widely used in past work (Amabile, 1982, Christensen et al., 1957, Kaufman et al., 2008 and Kaufman et al., 2007). For divergent thinking tasks, several trained raters simply evaluate and score individual ideas (Silvia, 2011, Silvia and Kimbrel, 2010 and Silvia et al., 2008) or the set of ideas (Silvia et al., 2009). Thus far, we have found that subjective ratings of creativity are unconfounded with fluency (Silvia et al., 2008) and that the relationships of creativity with intelligence are substantially larger (Nusbaum & Silvia, 2011a).

3. Cognitive abilities and metaphor production

Metaphor provides an interesting context for studying the role of intellectual abilities in creative cognition.1 How people generate metaphors is fascinating in its own right—despite the large literature on how people understand metaphor (Gibbs, 1994, Glucksberg, 2001 and Glucksberg et al., 1997), little is known about how people make metaphors. Creative metaphors are also good examples of real-world creativity, so metaphor provides a fruitful context for studying creative thought. Unlike divergent thinking, which many critics contend is unrealistic and artificial (Sawyer, 2006 and Simonton, 1999), metaphors are a common and valued form of creativity in speech and writing (Plotnik, 2007).

We propose that producing creative metaphors, like producing creative responses to divergent thinking tasks, involves several executive processes. The mechanics of metaphor production are just beginning to receive attention (see Chiappe and Chiappe, 2007 and Pierce and Chiappe, 2009), but models of metaphor comprehension provide insight into how people might compose metaphors. In the property attribution model of metaphor (Glucksberg, 2001 and Glucksberg et al., 1997), metaphors entail attributing a property of a vehicle to a topic. In the metaphor “Some toddlers are tyrants,” for example, the “demanding and domineering” feature of the vehicle (“tyrants”) is attributed to the topic (“some toddlers”). To understand the metaphor, people create a superordinate “attributive category” (“things that are demanding and domineering”) that the vehicle exemplifies and that can plausibly include the topic.

Using the property attribution model as a guide, we can see how creating a metaphor involves several executive processes. First, people must choose a property that they wish to attribute to the topic. For the topic “teaching,” for example, people must select what they wish to say about teaching (e.g., that it is rewarding, stressful, challenging, or unpredictable). Second, people must then scan semantic knowledge for suitable vehicles that exemplify the abstract, higher-order attributive category (e.g., searching for “things that are stressful”). Doing so requires maintaining access to the category while inhibiting many kinds of knowledge: features of the topic and of possible vehicles that are irrelevant to the higher-order category (cf. Gernsbacher, Keysar, Robertson, & Werner, 2001); highly accessible but irrelevant semantic knowledge (e.g., adjectival descriptions of the topic); and the many accessible but trite possibilities, such as idioms, clichés, and dead metaphors. Finally, likely vehicles (e.g., "lion taming") must be evaluated according to abstract criteria (e.g., "Does this metaphor convey the desired meaning and emotional tone? Is it clever or interesting?"), revised, and then retained or discarded.

Consistent with our analysis, the small body of work on how people make metaphors suggests that several cognitive abilities—including executive abilities—are involved. Taylor (1947) conducted one of the earliest studies of cognitive abilities and metaphor production. He developed a similes task that presented incomplete metaphor stems (e.g., "His skin was as brown as _____") and required participants to complete the stem three different ways. The similes task loaded on ideational fluency and verbal versatility factors. Interestingly, Taylor suggested an executive mechanism for the verbal versatility factor (p. 251):

"a person who is good in this ability can readily break the set of the first answer and produce a second answer, and then a third answer, that expresses the same general meaning. Others may find it difficult to break away from the first answer to restate the same idea in a somewhat different form."

Consistent with a role for interference management, the similes task had moderate correlations ($r = .32$ and $r = .37$) with measures of inductive reasoning.

Guilford and his research group developed several metaphor completion tasks (e.g., simile insertion and simile completion) as part of their research on verbal fluency (Christensen and Guilford, 1963 and Merrifield et al., 1963). Similar tests appear in the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). Contemporary research has extended this assessment approach. In a recent study (De Barros, Primi, Miguel, Almeida, & Oliveira, 2010), people were asked to complete nine metaphors (e.g., "A camel is the _____ of the desert") with up to four responses. Metaphor scores (primarily the number and quality of the responses) covaried with measures of analogical reasoning.

A different method for assessing metaphor production—the figurative statement production task (Pierce & Chiappe, 2009)—presents people with a topic (e.g., "Some jobs are _____") as well as a property to be attributed to the topic (e.g., “Confining and constraining, and make you feel like you are just putting in time”). People must produce a metaphor by generating a vehicle that successfully attributes the property to the topic (e.g., jails, prisons). In several studies, Chiappe and his colleagues (Chiappe & Chiappe, 2007, Studies 2 and 3; Pierce & Chiappe, 2009) have shown that working memory span—a cognitive ability central to reasoning and executive abilities (e.g., de Abreu et al., 2010, Kane et al., 2004 and Süß et al., 2002)—influences how people make metaphors. People with larger working memory spans generated better metaphors, and the effects of working memory were significant after controlling for individual differences in verbal fluency and vocabulary knowledge.

4. The present research

In the present research, we examined the contribution of fluid intelligence (Gf)—the ability associated with using “deliberate and controlled mental operations to solve novel problems that cannot be performed automatically” (McGrew, 2009, p. 5)—to the generation of creative metaphors. The sentence-completion tasks used in past work assess how people generate conventional metaphors—responses that are apt, familiar, and easy to understand, such as “Some lawyers are sharks” and “Some jobs are jails.” It's obviously important to understand conventional metaphors, but to assess how people generate creative metaphors—responses that are novel, original, and unfamiliar—different methods are needed. People can get high scores on sentence-completion measures by retrieving accessible and obvious vehicles, and most people will give the same handful of responses. To generate creative metaphors, in contrast, people need to inhibit such obvious responses and instead develop a response that few other people will give. It's thus likely that conventional and creative metaphors rely on different abilities, an issue we revisit in the General Discussion.

To assess creative metaphor production, we asked people to generate two metaphors that described past emotional experiences. We emphasized that people should try to come up with a metaphor that was creative, clever, unique, and interesting rather than a conventional expression, dead metaphor, or standard idiom. These metaphors were rated for creativity by three raters who applied methods for subjective scoring that were developed and refined in other domains of creativity research (e.g., Kaufman et al., 2008 and Silvia et al., 2008). To measure Gf, we administered six tests of reasoning, primarily inductive reasoning (Carroll, 1993). All six tasks were essentially visual and spatial. By omitting Gf tasks with significant verbal aspects (e.g., verbal reasoning tasks), we can reduce superficial overlap between the Gf and metaphor tasks and thus provide a stricter test of their relationship.

We also sought evidence for incremental validity of Gf, particularly with regards to personality. Some personality traits predict both intelligence and creativity, so they are potential “third variables” that should be examined (Silvia, 2008a). The most notable trait is openness to experience. Of the many personality traits that predict creativity, openness to experience has the largest and most consistent effects across a range of samples and tasks (Feist, 1998 and Silvia et al., 2009b). Moreover, of the five broad factors, openness to experience has the largest relationship with intelligence (Ashton et al., 2000, DeYoung, 2011 and Nusbaum and Silvia, 2011b). We thus evaluated openness to experience as a possible “third variable” that might explain a relationship between Gf and the generation of creative metaphors.

5. Method

5.1. Participants

The sample was comprised of 132 undergraduates—91 women, 41 men—at the University of North Carolina at Greensboro. Students volunteered to participate and received credit toward a research option in a psychology class. One person skipped past both metaphor tasks and was excluded, leaving a final sample of 131 people. The self-identified racial and ethnic composition of the sample was primarily European-American (61%), African American (26%), Asian American (10%), and Hispanic/Latino (5%). (Participants could select more than one category, and approximately 6% did so.)

5.2. Procedure

The study was carried out in a group setting, with the number of participants ranging from 1 to 8. Upon entering the lab room, participants were given a consent form and a brief explanation of the study procedures. Following informed consent, students completed a series of fluid intelligence tests, a metaphor generation task, and some self-report questionnaires. The tasks and questionnaires were administered with MediaLab v2010.

5.2.1. Fluid intelligence (Gf) tasks

Participants completed six measures of fluid intelligence (Gf). Most of the tasks focused on inductive reasoning, and all of them were primarily non-verbal or spatial: (1) odd-numbered items from the Ravens Advanced Progressive Matrices (18 items, 12 min); (2) a letter sets task (16 items, 4 min), in which people must decide which set of four letters violates a rule followed by the others (Ekstrom et al., 1976); (3) a number series task (15 items, 4.5 min), in which people must discern the rule governing a string of numbers to choose the correct number in the sequence (Thurstone, 1938); (4) the series task (13 items, 3 min) from the Culture Fair Intelligence Test (CFIT; Cattell & Cattell, 1961/2008), which involves choosing an image that correctly completes a series of images; (5) the matrices task from the CFIT (13 items, 3 min), which involves deciding which item completes the pattern in a matrix; and (6) a paper folding task (10 items, 3 min), in which people indicate what a piece of paper would look like after being folded, punched with holes, and unfolded (Ekstrom et al., 1976).

5.2.2. Metaphor production task

After completing the Gf tasks, the participants were asked to come up with two metaphors that described personal experiences. The experimenter first briefly defined and explained what metaphors were and gave examples of three common kinds of metaphor structures: metaphors (e.g., “All the world is a stage”), similes (e.g., “Justice is like a train that is nearly always late”), and compound metaphors (e.g., “Life is like a box of chocolates: you never know what you're going to get”). People were then given a prompt and asked to come up with a metaphor. We used two prompts. For the first metaphor, people were told “Think of the most boring high-school or college class that you've ever had. What was it like to sit through?”; for the second metaphor, people were told “Think about the most disgusting thing you ever ate or drank. What was it like to eat or drink it?” To help get them started, we provided several examples of stems for the first metaphor (e.g., “Being in that class was like…,” “That class was…”). People could spend as much time as they wished composing each metaphor—the software recorded the amount of time. Past work has found that intense autobiographical emotional experiences elicit relatively higher rates of novel metaphors (Fainsilber and Ortony, 1987 and Williams-Whitney et al., 1992).

As in our divergent thinking research (Nusbaum and Silvia, 2011a and Silvia et al., 2008), we instructed people to “be creative.” Many studies have shown that measures of creativity are more valid when people are trying to generate creative responses (e.g., Harrington, 1975 and Niu and Liu, 2009). The participants were told that “the aim is to come up with something creative—something clever, humorous, original, compelling, or interesting.”

Three raters judged each metaphor independently: they were unaware of the other raters' scores and all information about the participants, including the participants' other metaphor. As a group, the raters had relatively high expertise and experience. They were graduate students conducting research on the psychology of creativity, and they had previous experience applying subjective scoring methods to divergent thinking tasks. For each prompt, the metaphors were identified by a random number and then sorted alphabetically. The raters scored each metaphor on a 5-point scale, anchored by 1 (not at all creative) and 5 (very creative). The single holistic score was based on several factors taken from our work on divergent-thinking scoring (Silvia et al., 2008): novelty (Was the metaphor original? Was it merely a cliché or a dead metaphor?), remoteness (Was the vehicle conceptually distant?), and cleverness (Was the metaphor interesting, funny, striking, or incisive?). Metaphors that received low scores tended to be common idioms that people retrieved from memory. For example, many people said that sitting through a boring class was like “watching paint dry” or “watching grass grow.” Metaphors that received higher scores tended to be original, clever, and elaborated, such as “Trying to stay awake during that class was like trying not to get seconds at an all-you-can-eat buffet” and “Eating escargot is the picky eater's death penalty.”

5.2.3. Questionnaires

After the metaphor task, people completed demographic items as well as the personality scales. We used the 60-item NEO Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992) to assess the five major factors of personality (McCrae & Costa, 1997), including openness to experience. Each of the five factors is measured with 12 items, and people respond to each item using a five-point scale (1 = strongly disagree, 5 = strongly agree).

6. Results

6.1. Data reduction and modeling

We examined the relationship between Gf and metaphor creativity using structural equation modeling. Gf was specified as a latent variable with six indicators (see Fig. 1). Each Gf task served as an indicator. The indicators were centered, and the variance of the latent factor was fixed to 1. A confirmatory factor analysis (CFA) of this measurement model showed good fit: $\chi^2(9\ df) = 14.19$, $p = .12$, CFI = .955, SRMR = .047, RMSEA = .066 (90% CI = .00, .13). The reliability of a latent variable can be estimated with coefficient *H*, known as *construct reliability* or *maximal reliability* (Drewes, 2000, Hancock and Mueller, 2001 and Silvia, 2011). *H* ranges from 0 to 1 and expresses the “proportion of variability in the construct explainable by its own indicator variables” (Hancock & Mueller, 2001, pp. 202–203). For the Gf CFA, *H* was .76.

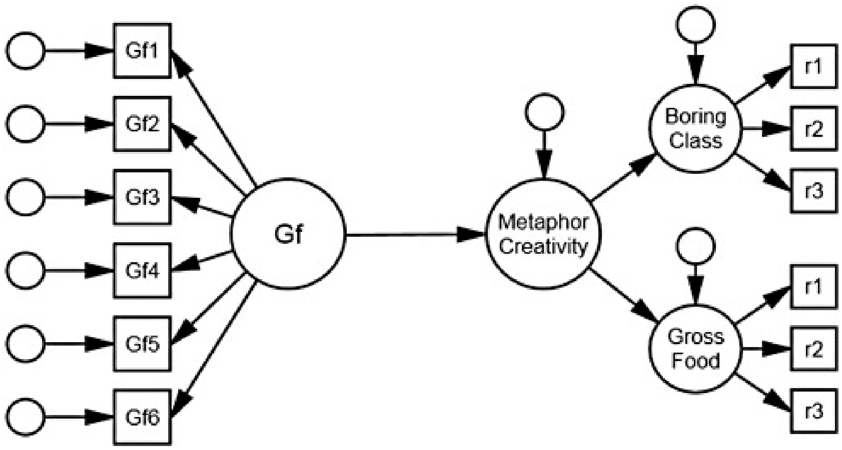

Fig. 1. A depiction of the latent variable model for fluid intelligence and metaphor creativity. The ratings of the metaphor tasks (r1, r2, and r3) are ordinal and thus lack residual variances.

Metaphor quality was specified as a higher-order latent variable defined by two lower-order latent variables: the “boring class” metaphor and the “disgusting food” metaphor. For identification, the two paths were constrained to be equal. In turn, each lower-order latent variable had each rater's scores as the three indictors (see Fig. 1). The variances for the latent variables were fixed to 1. The raters' scores were highly skewed, which is typical for ratings of creative products. We thus modeled the ratings as ordinal variables, which avoids violating the assumption of multivariate normal indicators (Kline, 2010, Skrondal and Rabe-Hesketh, 2004 and Yang-Wallentin et al., 2010). An ordinal CFA has the virtue of properly modeling the indicators, but one limitation is that it doesn't afford the conventional indices of model fit.2 Reliability was good for both the “boring class” ratings (H = .86) and “disgusting food” ratings (H = .67).

Table 1 displays descriptive statistics and correlations. All analyses were conducted with Mplus 6.1 using maximum likelihood estimation with robust standard errors. All regression coefficients are standardized.

Table 1. Descriptive statistics and correlations.

	M	*SD*	Min, Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Ravens	9.99	3.25	0, 16	1																		
2. Paper folding	5.23	2.28	1, 10	.50	1																	
3. Series completion	7.93	1.52	4, 11	.42	.42	1																
4. Matrix completion	6.46	1.05	4, 9	.19	.31	.28	1															
5. Letter sets	8.79	2.48	3, 14	.31	.33	.29	.20	1														
6. Number series	8.03	1.99	4, 14	.31	.25	.24	.10	.42	1													
7. Neuroticism	2.82	.66	1.00, 4.58	.05	.02	− .12	− .09	.16	.05	1												
8. Extraversion	3.57	.52	1.83, 4.67	− .09	− .16	− .12	− .14	− .07	.03	− .21	1											
9. Openness to experience	3.35	.51	2.17, 4.67	.23	.10	.16	− .06	.03	− .02	− .04	− .09	1										
10. Agreeableness	3.52	.48	2.17, 4.58	.01	.03	.12	.11	.00	.04	− .22	.10	.06	1									
11. Conscientiousness	3.48	.50	2.00, 4.92	− .18	− .02	− .17	− .08	− .08	.00	− .26	.30	− .23	.17	1								
12. Boredom: rater 1	1.40	.74	1, 4	.23	.09	.10	.00	.04	.10	− .07	− .04	.32	− .06	− .09	1							
13. Boredom: rater 2	1.28	.60	1, 4	.27	.16	.13	− .01	.13	.10	.09	− .12	.09	.09	− .16	.34	1						
14. Boredom: rater 3	2.24	1.05	1, 4	.19	.17	.20	− .05	.17	.03	.11	− .21	.21	− .08	− .26	.54	.27	1					
15. Disgust: rater 1	1.29	.56	1, 3	.20	.12	.08	− .07	.04	.14	− .02	− .06	.16	.15	.01	.15	.17	.09	1				
16. Disgust: rater 2	1.35	.69	1, 4	.12	.12	− .01	.04	.03	.05	.10	− .11	.06	.08	− .07	.13	.09	.20	.19	1			

	M	*SD*	Min, Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
17. Disgust: rater 3	1.86	.85	1, 4	.16	.10	.00	.07	.11	.21	.20	− .06	.00	− .09	− .13	.12	.16	.20	.28	.36	1		
18. Boredom time	105.75	73.72	17.30, 556.29	.16	.18	.09	.01	.06	− .01	− .10	− .17	.18	.19	− .11	.25	.30	.20	.28	.14	.17	1	
19. Disgust time	96.72	76.41	16.66, 632.33	.10	.09	− .04	− .09	− .13	− .09	− .03	− .32	.22	.09	− .09	.19	.34	.28	.15	.15	.16	.47	1

Note. *n* = 131. The time scores are in seconds. The latent variable models treated the raters' scores as ordinal variables (see text for details), so analyses based on the covariance matrix will vary slightly. Researchers interested in reanalyzing the data can obtain the raw data and Mplus input files from the first author.

6.2. Role of Gf in metaphor creativity

Did Gf predict how well people came up with creative metaphors? Our first model, shown in Fig. 1, estimated the effect of Gf on metaphor creativity. Gf had a significant effect, β = .49, p = .011 (see Table 2). Furthermore, the effect size was large, using the benchmarks of .10 for small, .30 for medium, and .50 for large (Cohen, 1988). Gf explained 24.3% of the variance in metaphor creativity scores.

Table 2. Summary of the regression effects.

Model	Predictor	Standardized beta	Standard error	*p*value	95% Confidence interval
1. Gf only	Gf	.49	.19	.011	.11, .87
2. Gf and personality	Gf	.41	.20	.038	.02, .79
	Neuroticism	.14	.18	.435	− .21, .49
	Extraversion	− .07	.14	.594	− .35, .20
	Openness to experience	.26	.14	.063	− .01, .53
	Agreeableness	− .05	.15	.742	− .34, .24
	Conscientiousness	− .18	.16	.256	− .50, .13

Note. *n* = 131.

6.3. Incremental validity

Did Gf continue to predict metaphor creativity after controlling for personality? We estimated a model in which Gf and the five factors of personality were predictors of metaphor creativity. The five factors were modeled as observed variables. As shown in Table 2, Gf continued to have an effect that was both significant and medium-to-large in size, β = .41, *p* = .038. None of the five factors had a significant effect, but openness to experience had a marginal effect. The largest effects were associated with openness to experience (β = .26, *p* = .063) and conscientiousness (β = − .18, *p* = .256). Consistent with the large literature on personality and creativity, openness predicted higher creativity and conscientiousness predicted lower creativity (Batey and

Furnham, 2006 and Feist, 1998). As a set, Gf and the five factors explained 34.8% of the variance in metaphor creativity scores.

6.4. Additional descriptive analyses

People could spend as much time as they wished on each metaphor, so how long did they spend? Table 1displays the descriptive statistics for time (in seconds). In general, people spent about a minute and a half on each metaphor, but the variability was large. As Table 1 shows, time ranged from around 17 s to around 10 min per metaphor, and people's time spent writing the first metaphor correlated significantly with time spent on the second ($r = .47$, $p < .001$).

To examine the relationship of time with metaphor creativity, we formed a latent variable from the two time scores and constrained the two indicator paths to be identical. The latent time variable was highly correlated with metaphor creativity ($r = .70$, $p < .001$)–the worst metaphors were generated the fastest. Time was only modestly related to Gf ($r = .14$, $p = .282$), however, so the effect of Gf on metaphor could not merely be due to spending more time on the tasks. Instead, time was significantly predicted by several factors of personality. A regression model with the five factors as predictors and time as the outcome found significant effects for extraversion ($\beta = -.37$, $p < .001$), openness to experience ($\beta = .23$, $p = .009$), and agreeableness ($\beta = .20$,$p = .012$).

Finally, we explored the role of gender in metaphor creativity. In our sample, men received significantly higher scores on the metaphor task, $\beta = -.70$, $p = .031$. Because gender is a binary predictor, the coefficient is *Y*-standardized: it represents the difference in metaphor quality, in standard deviation units, between the two groups (Long, 1997).

7. General discussion

People high in fluid intelligence (Gf) made metaphors that were much more creative. The size of the effect ($\beta = .49$) is particularly notable in light of the lack of surface overlap between the domains—the Gf tasks were essentially non-verbal—and the cultural stereotype of metaphors as literary devices that would ostensibly not have much to do with reasoning abilities. From our perspective on intelligence and creative cognition, however, one would expect substantial contributions of fluid and executive abilities (Nusbaum & Silvia, 2011a). Producing an original metaphor involves choosing an abstract property to attribute, searching knowledge for a vehicle that has the abstract property, and evaluating and revising the resulting metaphor. Throughout, there is substantial interference from metaphor-irrelevant features of both the topic and vehicle (Gernsbacher & Robertson, 1999) as well as from highly accessible idioms, dead metaphors, and adjectives. Searching knowledge based on an abstract criterion and managing interference both entail executive processes (Unsworth, 2010), so we would expect a significant contribution of Gf to the quality of creative metaphors.

The empirical debate about intelligence and creativity has taken place, for the most part, within the domain of divergent thinking. Divergent thinking is central to creativity, in our opinion, because many real-world creative problems involve generating many ideas and then reworking them. At the same time, many critics have argued that the tasks are artificial (Sawyer, 2006). As Simonton (1999) put it, “creativity in music, for example, is not going to be very predictable on the basis of how many uses one can imagine for a toothpick” (p. 314). This perhaps misses the point—in cognitive research, artificial tasks are useful for understanding how mental processes operate—but it is nevertheless important to understand how people generate realistic creative products.

Unlike divergent thinking, metaphor is an uncontroversial example of real-world creativity: it is valued in writing and speaking, and many books on creative writing try to teach people how to generate creative and interesting metaphors (e.g., Plotnik, 2007). Our metaphor production task also varied from divergent thinking tasks in several ways: divergent thinking tasks are usually timed (cf. Wallach & Kogan, 1965) and encourage people to generate many brief ideas, whereas the metaphor task is untimed and encourages people to generate a single elaborated idea. By showing a substantial effect in a different creative domain, the present work further suggests that the long-standing notion that creativity and intelligence are independent abilities should be revisited (Nusbaum & Silvia, 2011a).

The exploratory analyses of time shed some light on how people generate metaphors. First, the fact that people who took longer came up with better metaphors is consistent with the kinds of cognitive processes that are employed. People with brief response times appeared to use a memory-based retrieval strategy (Gilhooly et al., 2007), in which they searched memory for an appropriate metaphor and then wrote it. Because this strategy yields things people have seen and heard before, it typically yields common idioms and clichés. People who took their time, in contrast, appeared to be developing an answer on-the-spot, which takes more time but yields a more original result. Second, the relationships between time and personality suggest different ways of engaging with the task. People high in extraversion took less time, consistent with extraversion's impulsive quality. It would be interesting to know if people high in extraversion spent less time generating an idea, revising and evaluating the idea, or both. In contrast, people high in openness to experience spent more time. Research shows that people high in openness to experience enjoy opportunities to use their imaginations and to be creative (Joy, 2001 and Joy, 2005), facts that fit with choosing to spend more time developing clever metaphors.

To extend the present work, future research on metaphor production should examine a broader range of factors within the Cattell–Horn–Carroll (CHC) model (Carroll, 1993 and McGrew, 2005). In particular, crystallized intelligence (Gc) and broad retrieval ability (Gr) seem like important factors. Gc, whether viewed as verbal ability or as acquired knowledge (Kan, Kievit, Dolan, & van der Mass, 2011), should be important to metaphor generation, given the highly verbal nature of the task. Moreover, Gr would seem to be important, given that creating a good metaphor involves scanning and manipulating stored semantic knowledge, from simple vocabulary to conceptual knowledge of possible topics and vehicles. Some studies have found an important role for Gc in creative thought (e.g., Greengross et al., 2012 and Sligh et al., 2005), but for the most part research on intelligence and creativity hasn't examined the range of CHC cognitive abilities. Including Gf, Gc, and Gr in the same study could clarify their unique contributions to creative metaphor production, and it would allow the estimation of a higher-order g factor and its effects.

Another fruitful direction for future research is to compare the role of cognitive abilities in creative and conventional metaphors. Most research on metaphor production prompts people to generate brief metaphors that are highly conventional and apt, whereas our task encourages people to generate an unconventional metaphor. It's likely that the cognitive processes and abilities differ across these tasks. People can successfully produce conventional metaphors (e.g., “Some jobs are jails”) by retrieving accessible semantic knowledge, including obvious tropes, clichés, and idioms. For creative metaphors, in contrast, people must usually inhibit the most obvious instances of a feature as well as accessible idioms and dead metaphors. Factors such as Gc and Gr may thus be relatively more important for conventional metaphors, whereas factors such as Gf may be relatively more important for creative metaphors.

8. Conclusion

Unconventional and clever metaphors are among the most creative uses of language, but the study of creative metaphor has received little attention. By showing strong effects of fluid intelligence on metaphor creativity, the present research extends the study of intelligence in two directions: it shows that Gf strongly predicts creativity in a domain other than divergent thinking, and it suggests that the Cattell–Horn–Carroll model of cognitive abilities could be fertile for understanding cognitive aspects of metaphor production.

Acknowledgments

We thank Emily Nusbaum and Bridget Smeekens for their help with this research.

References

T.M. Amabile. Social psychology of creativity: A consensual assessment technique. Journal of Personality and Social Psychology, 43 (1982), pp. 997–1013

M.C. Ashton, K. Lee, P.A. Vernon, K.L. Jang. Fluid intelligence, crystallized intelligence, and the openness/intellect factor. Journal of Research in Personality, 34 (2000), pp. 198–207

J.A. Barnden. Metaphor and metonymy: Making their connections more slippery. Cognitive Linguistics, 21 (2010), pp. 1–34

M. Batey, T. Chamorro-Premuzic, A. Furnham. Intelligence and personality as predictors of divergent thinking: The role of general, fluid and crystallised intelligence. Thinking Skills and Creativity, 4 (2009), pp. 60–69

M. Batey, A. Furnham. Creativity, intelligence, and personality: A critical review of the scattered literature. Genetic, Social, and General Psychology Monographs, 132 (2006), pp. 355–429

J.B. Carroll. Human cognitive abilities: A survey of factor-analytic studies. Cambridge University Press, New York (1993)

R.B. Cattell, A.K.S. Cattell. Measuring intelligence with the Culture Fair Tests. Hogrefe, Oxford, UK (1961/2008)

D.L. Chiappe, P. Chiappe. The role of working memory in metaphor production and comprehension. Journal of Memory and Language, 56 (2007), pp. 172–188

P.R. Christensen, J.P. Guilford. An experimental study of verbal fluency factors. British Journal of Statistical Psychology, 16 (1963), pp. 1–26

P.R. Christensen, J.P. Guilford, R.C. Wilson. Relations of creative responses to working time and instructions. Journal of Experimental Psychology, 53 (1957), pp. 82–88

J. Cohen. Statistical power analysis for the behavioral sciences. (2nd ed.)Lawrence Erlbaum Associates, Mahwah, NJ (1988)

P.T. Costa Jr., R.R. McCrae. Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Psychological Assessment Resources, Odessa, FL (1992)

P. de Abreu, A.A. Conway, S.E. Gathercole. Working memory and fluid intelligence in young children. Intelligence, 38 (2010), pp. 552–561

D.P. De Barros, R. Primi, F.K. Miguel, L.S. Almeida, E.P. Oliveira. Metaphor creation: A measure of creativity or intelligence? European Journal of Education and Psychology, 3 (2010), pp. 103–115

C.G. DeYoung. Intelligence and personality. R.J. Sternberg, S.B. Kaufman (Eds.), The Cambridge handbook of intelligence, Cambridge University Press, New York (2011), pp. 711–737

D.W. Drewes. Beyond the Spearman-Brown: A structural approach to maximal reliability. Psychological Methods, 5 (2000), pp. 214–227

R.B. Ekstrom, J.W. French, H.H. Harman, D. Dermen. Manual for Kit of Factor-Referenced Cognitive Tests. Educational Testing Service, Princeton, NJ (1976)

L. Fainsilber, A. Ortony. Metaphorical uses of language in the expression of emotions. Metaphor and Symbolic Activity, 2 (1987), pp. 239–250

G.J. Feist. A meta-analysis of personality in scientific and artistic creativity. Personality and Social Psychology Review, 2 (1998), pp. 290–309

R.A. Finke, T.B. Ward, S.M. Smith. Creative cognition: Theory, research, and applications. MIT Press, Cambridge, MA (1992)

L. Gabora. Creative thought as a non-Darwinian evolutionary process. Journal of Creative Behavior, 39 (2005), pp. 262–283

M.A. Gernsbacher, B. Keysar, R.R. Robertson, N.K. Werner. The role of suppression and enhancement in understanding metaphors. Journal of Memory and Language, 45 (2001), pp. 433–450

M.A. Gernsbacher, R.W. Robertson. The role of suppression in figurative language comprehension. Journal of Pragmatics, 31 (1999), pp. 1619–1630

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

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

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

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

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

G. Greengross, R.A. Martin, G. Miller. Personality traits, intelligence, humor styles, and humor production ability of professional stand-up comedians compared to college students. Psychology of Aesthetics, Creativity, and the Arts, 6 (2012), pp. 74–82

G.R. Hancock, R.O. Mueller. Rethinking construct reliability within latent variable systems. R. Cudeck, S. du Toit, D. Sörbom (Eds.), Structural equation modeling: Present and future, Scientific Software International, Lincolnwood, IL (2001), pp. 195–216

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

S. Joy. The need to be different predicts divergent production: Toward a social learning model of originality. Journal of Creative Behavior, 35 (2001), pp. 51–64

S.P. Joy. Innovation motivation and artistic creativity. Journal of Creative Behavior, 39 (2005), pp. 35–56

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

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

J.C. Kaufman. Creativity 101. Springer, New York (2009)

Kaufman et al., 2008. J.C. Kaufman, J. Baer, J.C. Cole, J.D. Sexton. A comparison of expert and nonexpert raters using the consensual assessment technique. Creativity Research Journal, 20 (2008), pp. 171–178

J.C. Kaufman, J. Lee, J. Baer, S. Lee. Captions, consistency, creativity, and the consensual assessment technique: New evidence of reliability. Thinking Skills and Creativity, 2 (2007), pp. 96–106

J.C. Kaufman, J.A. Plucker. Intelligence and creativity. R.J. Sternberg, S. Kaufman (Eds.), The Cambridge handbook of intelligence, Cambridge University Press, New York, NY (2011), pp. 771–783

K.H. Kim. Can only intelligent people be creative? Journal of Secondary Gifted Education, 16 (2005), pp. 57–66

K.H. Kim, B. Cramond, J. VanTassel-Baska. The relationship between creativity and intelligence. J.C. Kaufman, R.J. Sternberg (Eds.), The Cambridge handbook of creativity, Cambridge University Press, New York, NY (2010), pp. 395–412

R.B. Kline. Principles and practice of structural equation modeling. (3rd ed.)Guilford, New York (2010)

J.S. Long. Regression models for categorical and limited dependent variables. Sage, Thousand Oaks, CA (1997)

R.R. McCrae, P.T. Costa Jr. Personality trait structure as a human universal. American Psychologist, 52 (1997), pp. 509–516

K.S. McGrew. The Cattell–Horn–Carroll theory of cognitive abilities. D.P. Flanagan, P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (2nd ed.), Guilford, New York, NY (2005), pp. 136–181

K.S. McGrew. CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. Intelligence, 37 (2009), pp. 1–10

P.R. Merrifield, J.P. Guilford, P.R. Christensen, J.W. Frick. Interrelationships between certain abilities and certain traits of motivation and temperament. The Journal of General Psychology, 65 (1963), pp. 57–74

W. Niu, D. Liu. 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 (2009), pp. 93–98

E.C. Nusbaum, P.J. Silvia. Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. Intelligence, 39 (2011), pp. 36–45

E.C. Nusbaum, P.J. Silvia. Are openness and intellect distinct aspects of openness to experience? A test of the O/I model. Personality and Individual Differences, 51 (2011), pp. 571–574

K.U. Panther, L.L. Thornburg. Metonymy. D. Geeraerts, H. Cuyckens (Eds.), The Oxford handbook of cognitive linguistics, Oxford University Press, New York (2007), pp. 236–263

R.S. Pierce, D.L. Chiappe. The roles of aptness, conventionality, and working memory in the production of metaphors and similes. Metaphor and Symbol, 24 (2009), pp. 1–19

A. Plotnik. Spunk and bite: A writer's guide to bold, contemporary style. Random House Reference, New York (2007)

J.A. Plucker, M. Qian, S. Wang. Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. Journal of Creative Behavior, 45 (2011), pp. 1–22

F. Preckel, H. Holling, M. Wiese. Relationship of intelligence and creativity in gifted and non-gifted students: An investigation of threshold theory. Personality and Individual Differences, 40 (2006), pp. 159–170

F. Preckel, C. Wermer, F.M. Spinath. The interrelationship between speeded and unspeeded divergent thinking and reasoning, and the role of mental speed. Intelligence, 39 (2011), pp. 378–388

M.A. Runco. Creativity. Academic Press, San Diego, CA (2007)

R.K. Sawyer. Explaining creativity: The science of human innovation. Oxford University Press, New York (2006)

P.J. Silvia. Another look at creativity and intelligence: Exploring higher-order models and probable confounds. Personality and Individual Differences, 44 (2008), pp. 1012–1021

P.J. Silvia. Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan (1965). Creativity Research Journal, 20 (2008), pp. 34–39

P.J. Silvia. Subjective scoring of divergent thinking: Examining the reliability of unusual uses, instances, and consequences tasks. Thinking Skills and Creativity, 6 (2011), pp. 24–30

P.J. Silvia, N.A. Kimbrel. A dimensional analysis of creativity and mental illness: Do anxiety and depression symptoms predict creative cognition, creative accomplishments, and creative self-concepts? Psychology of aesthetics, creativity, and the arts, 4 (2010), pp. 2–10

P.J. Silvia, C. Martin, E.C. Nusbaum. A snapshot of creativity: Evaluating a quick and simple method for assessing divergent thinking. Thinking Skills and Creativity, 4 (2009), pp. 79–85

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

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

D.K. Simonton. Creativity as blind variation and selective retention: Is the creative process Darwinian? Psychological Inquiry, 10 (1999), pp. 309–328

A. Skrondal, S. Rabe-Hesketh. Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models. Chapman & Hall/CRC, Boca Raton, FL (2004)

A.C. Sligh, F.A. Conners, B. Roskos-Ewoldsen. Relation of creativity to fluid and crystallized intelligence. Journal of Creative Behavior, 39 (2005), pp. 123–136

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

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

L.L. Thurstone. Primary mental abilities. University of Chicago Press, Chicago, IL (1938)

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

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

O. Vartanian. Decision junctures in the creative process. O. Vartanian, D.R. Mandel (Eds.), Neuroscience of decision making, Psychology Press, New York, NY (2011), pp. 311–327

M.A. Wallach, N. Kogan. Modes of thinking in young children: A study of the creativity–intelligence distinction. Holt, Rinehart, & Winston, New York (1965)

D. Williams-Whitney, J.S. Mio, P. Whitney. Metaphor production in creative writing. Journal of Psycholinguistic Research, 21 (1992), pp. 497–509

F. Yang-Wallentin, K.G. Jöreskog, H. Luo. Confirmatory factor analysis of ordinal variables with misspecified models. Structural Equation Modeling, 17 (2010), pp. 392–423

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

D.L. Zabelina, M.D. Robinson, J.R. Council, K. Bresin. Patterning and non-patterning in creative cognition: Insights from performance in a random number generation task. Psychology of Aesthetics, Creativity, and the Arts, 6 (2012), pp. 137–145

1 Consistent with usage in cognitive linguistics (Barnden, 2010 and Grady, 2007), we use metaphor as a higher-order term that encompasses several kinds of figurative thought, such as metaphors, similes, and analogies. Although not alike in all respects, these share features that distinguish them from other classes of figurative thought, such as metonymy and irony (Gibbs, 1994 and Panther and Thornburg, 2007).

2 Although treating the ratings as categorical indicators is more appropriate, treating the indicators as continuous doesn't appreciably change our results. The fit of the model is good—$\chi2(52 \text{ df}) = 57.60$, $p = .28$, CFI = .976, SRMR = .054, RMSEA = .029 (90% CI = .004, .064)—and Gf has an identical effect on metaphor ($\beta = .49$, $p = .008$).