Individual Differences in Both Fluid and Crystalized Intelligence Predict Metaphor Comprehension

Dušan Stamenković (dusan.stamenkovic@filfak.ni.ac.rs)
Faculty of Philosophy & Center for Cognitive Sciences, University of Niš
Niš, Serbia

Keith J. Holyoak (holyoak@lifesci.ucla.edu)
Department of Psychology, University of California, Los Angeles
Los Angeles, CA, USA

Abstract

The nature of the mental processes involved in metaphor comprehension has been the focus of debate. Research related to this debate has mainly examined the comprehension of simple nominal metaphors. Here we take an individual-differences approach to examine the comprehension of slightly more complex metaphors, some taken from literary sources, using two types of comprehension tests (selecting an overall interpretation or else selecting a completion). In a series of metaphor-comprehension experiments with college students, we measured both fluid intelligence (using the non-verbal Raven's Progressive Matrices test) and crystalized verbal intelligence (using a new Semantic Similarities Test). Each measure had a dissociable predictive relationship to metaphor comprehension, at least for those of the more complex literary variety. The pattern of individual differences suggests that metaphor comprehension broadly depends on both crystalized and fluid intelligence, with the latter less important for relatively simple metaphors.

Keywords: metaphor, analogy, conceptual combination, similarity, individual differences, intelligence

Introduction

Metaphor is the use of language to describe one thing in terms of something else that is conceptually very different, as in the poet Theodore Roethke's lament, "my memory, my prison." Metaphor and related cognitive processes have been linked to creative thinking not only in poetry (Holyoak, 1982, in press), but also in many scientific fields (e.g., Dunbar & Klahr, 2012). In artificial intelligence, the goal of automatically detecting and comprehending metaphors encountered in text corpora represents a current frontier (e.g., Gagliano, Paul, Booten, & Hearst, 2016). Given its evident importance in human thinking and language, an important goal for cognitive science is to understand how people grasp metaphors.

Psychologists, linguists, and philosophers have advanced many alternative theories, but two general accounts of metaphor comprehension have been especially influential. One proposal has been that metaphor comprehension requires analogical reasoning to relate the target to the source.[1] The idea that metaphor is based on analogy originated with Aristotle, and was advanced in modern times by Black (1962). In psychology, the hypothesis was developed further by Tourangeau and Sternberg (1981, 1982), Trick and Katz (1986), and Gentner and Clement (1988). An alternative account, proposed by Glucksberg and Keysar (1992), claims that nominal metaphors are interpreted as categorization statements. On this view, when Roethke claims that his memory is a prison, the target (*memory*) is stated to be a member of a category specified by the source (*prison*), where the latter takes on an abstract meaning like "location of extended confinement," rather than its more specific meaning of a building that houses prisoners. Metaphor-as-categorization can be modeled as a kind of conceptual combination (Kintsch, 2000, 2001; Kintsch & Bowles, 2002).

Despite decades of research addressing the question of whether metaphor comprehension depends on analogy, categorization, or some mix of both (e.g., Bowdle & Gentner, 2005), no firm answer has emerged (for a recent review see Holyoak & Stamenković, 2017). Psychological studies have largely focused on simple nominal metaphors (e.g., "The lawyer is a shark"). It appears that these can generally be comprehended without involvement of the brain area most closely linked to complex analogical reasoning (rostrolateral prefrontal cortex; see meta-analyses by Bohrn, Altmann & Jacobs, 2012; Rapp, Mutschler, & Erb, 2012; Vartanian, 2012). Thus available neural evidence does not support the hypothesis that analogy plays a major role in comprehending simple metaphors (even when the metaphor is novel; see Cardillo et al., 2012). At the same time, even proponents of the categorization view have cautioned that not all metaphors can be comprehended on the basis of categorization (Glucksberg & Haught, 2006).

The present study was guided by three major aims. First, we wished to examine comprehension of metaphors that are somewhat more complex than the nominal form, including examples drawn from literary sources. Second, rather than continuing to focus on specific models of metaphor comprehension, we stepped back to consider the general role of metaphor as a bridge between human thinking and language. Classical theories of intelligence (Cattell, 1971) distinguish between *fluid* and *crystalized* intelligence, where fluid intelligence involves reasoning (often nonverbal) about novel problems detached from prior knowledge, and crystalized intelligence involves reasoning (typically verbal)

[1] In linguistics, the terms *topic* and *vehicle* are sometimes used for what we term the *target* and *source*. The latter terms are commonly used in discussions of analogical reasoning.

that draws upon prior knowledge. Metaphor comprehension seems likely to tap both of these basic forms of intelligence. Third, our focus on types of intelligence in turn led us to adopt an *individual-differences* approach to investigate metaphor comprehension.

A relatively small number of previous studies have investigated individual differences in cognitive factors that might impact processing of metaphors. Trick and Katz (1986) found positive correlations between people's scores on a test of analogical reasoning and ratings of the comprehensibility of metaphors, especially when the source and target were drawn from dissimilar categories. A measure of vocabulary knowledge (which would be expected to reflect crystalized intelligence) did not add any predictive power. Nippold and Sullivan (1987) reported that within a sample of children, perceptual analogical reasoning was related to verbal analogical reasoning, as well as to comprehension of proportional metaphors (albeit weakly). A measure of verbal analogical reasoning did not add any predictive power. Thus neither of these studies provided support for a role of crystalized verbal intelligence in metaphor comprehension.

Kazmerski, Blasko, and Dessalegn (2003) had their participants complete IQ and working-memory tests, and rate and interpret a set of metaphors. The IQ measure included both fluid and crystalized components. They found that low-IQ participants produced poorer-quality interpretations relative to high-IQ individuals. A vocabulary subtest predicted interpretation quality (in apparent contrast to the null finding reported by Trick & Katz, 1986). However, a measure of spatial working-memory did not correlate with verbal IQ and did not predict quality of metaphor interpretations (a finding apparently contrary to that reported by Nippold & Sullivan, 1987). Thus although overall IQ predicted quality of metaphor interpretations, Kazmerski et al.'s findings did not clearly distinguish the impact of fluid and crystalized intelligence.

In a study by Chiappe and Chiappe (2007), individuals who scored high on a working-memory test generated higher-quality interpretations of metaphors more quickly. Measures of inhibitory control (based on Stroop interference and intrusion errors on a memory test) also predicted metaphor processing (also see Pierce & Chiappe, 2008). Both working memory and inhibitory control are executive functions closely linked to fluid intelligence (Ackerman, Beier & Boyle, 2005). In a production task, Chiappe and Chiappe found that measures of vocabulary knowledge and exposure to print (linked to crystalized intelligence) also predicted metaphor quality. Indeed, the measures of crystalized intelligence yielded somewhat higher correlations with metaphor than did the measures of working memory.

Thus although findings have been mixed, at least the study by Chiappe and Chiappe (2007) suggests that both fluid and crystalized intelligence have an impact on metaphor interpretation and production. The present study sought additional evidence of potential individual differences in metaphor comprehension.

Experiment 1

In Experiment 1 we administered tests of both fluid and crystalized intelligence to participants who performed a task requiring comprehension of metaphors, selected from both literary and nonliterary sources.

Participants

A total of 76 undergraduate students at the University of California, Los Angeles (UCLA) (female = 50, male = 25, undeclared = 1; mean age = 21.1) participated in the study for course credit. They were either native speakers of English, or bilinguals who spoke English fluently (self-assessed). Data from an additional five participants were dropped from analyses based on criteria indicative of carelessness or inattention on the verbal tasks: score of 12 or lower on the Semantic Similarities Test (max = 40), or 5 or lower (max = 20) on each set of metaphors, or extremely short overall response time (under 15 minutes for the entire set of tasks).

Design, Materials, and Procedures

Participants completed three tasks in a fixed order.

Raven's Progressive Matrices (RPM). Scores on this non-verbal intelligence test (Raven, 1938) correlate highly with measures of working memory as well as analogical reasoning (Snow, Kyllonen & Marshalek, 1984). The RPM is generally considered to be a highly reliable measure of fluid intelligence; to the best of our knowledge, it has never been used previously in conjunction with a test of metaphor comprehension. We used a short form of this test (Arthur et al., 1999), adapted for computer administration using *SuperLab* software.

Semantic Similarities Test (SST). [2] We created a new instrument to provide a rapid assessment of crystalized verbal intelligence with face validity of relevance to metaphor comprehension. The SST is designed to measure participants' ability to identify similarities between concepts expressed as single words. The test comprises 20 items (pairs of words), ordered from easy-to-hard based on preliminary data. For each pair, participants answered the question, "How are the two concepts in each pair similar to one another?" The instructions included two examples (*chair-sofa* and *turtle-tank*, for which the answers provided were "both are types of furniture" and "both have a form of armor", respectively).

An answer key was compiled based on pilot testing, which allowed us to score the participants' responses as fully correct (2), partially correct (1) or incorrect (0). An example of an easy item is *bird-airplane* (correct answer: "flies", or "both have wings"). More difficult is *tavern-church* (correct answer: "public building" or "a place of gathering"). The most difficult pairs were taken from nominal metaphors. An

[2] The SST with scoring key is available from the authors upon request.

example of a difficult item is *love-drug* (correct answer: "addictive", or "affects brain/thinking").

The resulting scale had an acceptable level of internal consistency (Cronbach's alpha = 0.61, based on data from 280 participants). Whereas the RPM is a formal and nonverbal test in which semantic knowledge plays virtually no role, the SST is a verbal test in which semantic knowledge of word meanings is critical. The task is similar to the Wechsler Adult Intelligence Scale (WAIS) Similarities subscale (a measure of verbal comprehension), but with entirely different items. The RPM and SST thus complement each other as relatively pure measures of fluid and crystalized verbal intelligence, respectively. We would, however, expect scores on the two tests to be correlated, as both should load on the *g* factor (general intelligence; Ackerman et al., 2005; Spearman, 1927).

Metaphor comprehension. The final task in this experiment consisted of 40 metaphor comprehension items, 20 from literary sources and 20 nonliterary. The literary metaphorical statements were selected from a list of literary metaphors drawn from poetry anthologies by Katz et al. (1988). The metaphors we chose were rated high on a goodness scale in the Katz et al. study; e.g., "The tongue is a bayonet." Their syntactic forms included nominal (*X is Y*), nominal with an adjective modifier, and nominal with a prepositional phrase.

The nonliterary metaphors included 20 items, some of them adapted from word pairs generated by Green et al. (2010, 2012) to make proportional verbal analogy problems in the form *A:B* :: *C:D* (e.g., *roof:house* :: *hat:man*). By dropping the D term, we converted some of these items into proportional metaphors in the form *A is the C of B* (e.g., "A roof is the hat of a house"). We augmented the set with similar items that we created following the same pattern. The literary and nonliterary items were intermixed and presented in a randomized order.

Comprehension was assessed by a task requiring selection of the best interpretation. For each metaphorical statement, three potential interpretations were provided, and the participants were asked to select the correct one. Examples of the interpretation task, for both literary and nonliterary metaphors, are shown in Table 3. (The examples in Table 3 are drawn from metaphors used in Experiment 2, which overlapped with those in the sets tested in Experiment 1.)

The stimuli for all three tasks were presented on a computer screen and participant responses were recorded. Instructions for each task were given immediately preceding that task. There was no time limit on any task, but participants were instructed to complete each task as quickly as possible.

Results

Performance on each task is summarized in Table 1.

Table 1. Descriptive statistics for each test (Experiment 1).

Test	Mean	Max	SD	Range
RPM	6.64	12	2.90	0–12
SST	29.03	40	4.05	19–37
Literary metaphors	15.90	20	2.91	6–20
Nonliterary metaphors	18.49	20	2.33	7–20

Correlation and regression analyses were performed to assess the interrelationships among the RPM, SST and metaphor comprehension. RPM and SST scores were moderately correlated with each other ($r(76) = .31$, $p = .006$). As summarized in Table 2, each individual-difference measure was correlated with accuracy on the comprehension test for both literary and nonliterary metaphors (individual correlations ranging from .37 to .49, $p < .01$ in all cases). Multiple regression analyses revealed that for both types of metaphors, RPM and SST scores each predicted separable variance in comprehension accuracy, with partial correlations ranging from .26 (RPM for nonliterary metaphors, $p < .05$) to .42 (SST with nonliterary metaphors, $p < .001$). This pattern suggests that while both measures have an impact on metaphor comprehension, RPM (fluid intelligence) may be somewhat less important than SST (crystalized intelligence) for the nonliterary metaphors.

Table 2. Correlations and partial correlations of individual-difference measures with metaphor comprehension (Experiment 1).

	Literary		**Nonliterary**	
	Correlation	**Partial Correlation**	**Correlation**	**Partial Correlation**
RPM	.43***	.34**	.37**	.26*
SST	.43***	.34**	.49***	.42***

*** p <.001; ** p< .01; * p < .05

Experiment 2A

Experiment 2 was designed to extend the findings of Experiment 1 by using multiple assessments of metaphor comprehension. In addition to the interpretation task used in Experiment 1 (select the best interpretation from a set of options), we also used a completion task (select the best word to complete a metaphor from a set of options). To avoid repeating items with different tasks, the 2 x 2 design (literary/nonliterary metaphors x interpretation/completion task) was decomposed into two pairs of conditions, which were run and analyzed separately. Table 4 shows examples of each type of metaphor with each comprehension task. Experiment 2A examined literary metaphors with the completion task and nonliterary metaphors with the comprehension task; Experiment 2B examined literary metaphors with the interpretation task and nonliterary metaphors with the completion task. We will introduce all four conditions as we describe Experiment 2A.

Participants

A total of 101 undergraduate UCLA students (female = 77, male = 23, undeclared = 1; mean age = 20.1) participated in the study for course credit. They were either native speakers of English, or bilinguals who spoke English fluently (self-assessed). Data from an additional 11 participants were dropped from analyses based on criteria indicative of carelessness or inattention on the verbal tasks: score of 12 or lower on the Semantic Similarities Test (max = 40), or 4 or lower (max = 15) on each set of metaphors, or extremely short overall response time (under 15 minutes for all tasks).

Table 3. Examples of literary and nonliterary metaphors in completion and interpretation tasks (Experiment 2).

Literary Metaphor Comprehension		Nonliterary Metaphor Comprehension	
Completion	**Interpretation**	**Completion**	**Interpretation**
Water is the blood of soft _________. 1) dreams 2) snows* 3) air	**The expression *Water is the blood of soft snows* means:** 1) Water brings coldness. 2) Water originates from soft snows.* 3) Soft snows are thicker than water.	**An election is the __________ of votes.** 1) cultivation 2) sowing 3) harvest*	**The expression *An election is the harvest of votes* means:** 1) Candidates collect signatures in an election. 2) Elections are scheduled at the same time as harvests. 3) Candidates collect votes in an election.*
__________ is a leaf in the gardens of God. 1) Goddess 2) Man* 3) Mother	**The expression *Man is a leaf in the gardens of God* means:** 1) God cherishes human beings.* 2) God waters the soil. 3) Human beings love God.	**A tire is the __________ of a car.** 1) shoe* 2) ankle 3) elbow	**The expression *A tire is the shoe of a car* means:** 1) Tires and shoes have the same patterns. 2) Tires are made in the same way as shoes. 3) Tires help cars move on the ground.*
The _________ is a rope that binds heaven and earth. 1) mind 2) body 3) soul*	**The expression *The soul is a rope that binds heaven and earth* means:** 1) The soul contains both heaven and earth. 2) The soul is what makes heaven look like earth. 3) The soul allows one to travel from earth to heaven.*	**__________ is the morning of life.** 1) Old age 2) Adulthood 3) Childhood*	**The expression *Childhood is the morning of life* means:** 1) Childhood is initiated before life. 2) Childhood comes at the same time as life. 3) Childhood comes early in life.*

* indicates the response scored as correct

Design, Materials, and Procedures

As in Experiment 1, all participants completed the RPM, SST, and metaphor comprehension tasks, in that order. The metaphors used in Experiments 2A and 2B (15 literary and 15 nonliterary) were generally selected from among those used in Experiment 1, with a few revisions. We also revised some of the options, aiming to make them more challenging. Table 3 presents some examples.

In Experiment 2A, the comprehension task consisted of two conditions, administered in a fixed order. The first condition used literary metaphors with a completion task, in which each metaphor was presented with a blank (e.g., *Sunlight is a golden* ________.) for which a completion was to be chosen. Three options were presented underneath, one of which (scored as correct) was from the original metaphor (for this example, *dust*).

The second condition used 15 nonliterary metaphors with the task of choosing the best interpretation, as in Experiment 1. Within all metaphor-related tasks, the items were displayed in a randomized order for each participant.

For all tasks, stimuli were displayed on a computer screen and participant responses were recorded. Participants received the instructions for each task separately, just before the relevant task. There was no time limit for any task, but participants were instructed to complete the task as quickly as possible.

Results

Performance on each task is summarized in Table 4.

Table 4. Descriptive statistics for each test (Experiment 2).

A: Results of Experiment 2A

Test	Mean	Max	SD	Range
RPM	7.11	12	2.67	2–12
SST	30.02	40	3.83	15–37
Literary metaphors (completion)	10.09	15	1.88	5–15
Nonliterary metaphors (interpretation)	14.08	15	1.28	7–15

B: Results of Experiment 2B

Test	Mean	Max	SD	Range
RPM	6.55	12	2.27	1–11
SST	28.99	40	3.91	14–37
Literary metaphors (interpretation)	11.49	15	1.82	7–15
Nonliterary metaphors (completion)	12.02	15	1.66	7–15

As in Experiment 1, correlation and regression analyses were performed to assess the interrelationships among the RPM, SST and metaphor comprehension. RPM and SST scores were again reliably correlated with each other ($r(101) = .32$, $p = .001$). As summarized in Table 5, both individual-difference measures were correlated with accuracy on the completion task with literary metaphors. Partial correlations revealed that each of the two individual-difference measures contributed separately to predicting performance for this condition. However, for the simpler nonliterary metaphors used in the interpretation task, only SST scores were a reliable predictor of performance (based on both raw and partial correlations).

Table 5. Correlations and partial correlations of individual-difference measures with metaphor comprehension (Experiment 2).

A: Pattern in Experiment 2A

	Literary		Nonliterary	
	Correlation	**Partial Correlation**	**Correlation**	**Partial Correlation**
RPM	.38***	.30**	.19	.10
SST	.36***	.27**	.31**	.26**

B: Pattern in Experiment 2B

	Literary		Nonliterary	
	Correlation	**Partial Correlation**	**Correlation**	**Partial Correlation**
RPM	.34***	.27**	.17	.08
SST	.30**	.21*	.32**	.28**

*** p <.001; ** p< .01; * p < .05

Experiment 2B

Method

Experiment 2B was identical to Experiment 2A, except that the metaphor task consisted of the interpretation condition with nonliterary metaphors, followed by the completion condition with literary metaphors.

A total of 103 undergraduate UCLA students (female = 70, male = 33; mean age = 20.31) participated in the study for course credit. They were either native speakers of English, or bilinguals who spoke English fluently (self-assessed). Data from an additional nine participants were dropped from analyses based on the same criteria as in Experiment 2A.

Results

Performance on each task is summarized in Table 4B. Correlation and regression analyses were again performed to assess the interrelationships among the RPM, SST and metaphor comprehension. RPM and SST scores were again reliably correlated with each other ($r(103) = .31, p = .001$). As summarized in Table 5B, both individual-difference measures were correlated with accuracy on the interpretation task with literary metaphors. Partial correlations revealed that each of the two individual-difference measures contributed separately to predicting performance for this condition. For the simpler nonliterary metaphors used in the completion task, only SST yielded a significant raw correlation with metaphor performance, and also a reliable partial correlation.

Discussion

The present study took an individual-differences approach to examine the cognitive factors that impact comprehension of metaphors. Following Katz et al. (1988), we differentiated between metaphors that originated in literary sources and those from nonliterary sources (generally constructed by psychologists for experimental purposes). We used two tasks to assess comprehension (respectively requiring choice of the best interpretation, or the best completion, for each metaphor), and also had participants complete two tests that assess aspects of cognition that might modulate the ability to grasp metaphors: the RPM (a standard measure of fluid intelligence), and a new Semantic Similarities Test (SST). The SST, designed to tap into the kind of crystalized verbal intelligence that might be expected to impact metaphor comprehension, has greater face validity than previously-used verbal measures, such as vocabulary knowledge.

The two ability tests proved to be moderately correlated with one another, as would be expected given the evidence of a general (*g*) factor in intelligence. However, the pattern of correlations with metaphor comprehension differed between the two tests. In each of three experiments, scores on the SST contributed unique variance to prediction of comprehension accuracy for both literary and nonliterary metaphors. In contrast, RPM was a robust and separable predictor for literary metaphors, but its contribution to predicting performance with the simpler nonliterary metaphors was weaker (Experiment 1) or statistically undetectable (Experiments 2A and 2B).

The present findings are in accord with Chiappe and Chiappe's (2007) evidence that both fluid and crystalized intelligence affect metaphor comprehension, with crystalized intelligence being the more potent factor (at least for simpler metaphors). Given the strong association between the RPM and measures of analogical reasoning (Snow et al., 1984), the relative weakness of the connection between RPM scores and comprehension of simple metaphors casts further doubt on the hypothesis that complex analogical reasoning is necessary to understand such metaphors (Holyoak & Stamenković, 2017).

Our findings are, however, consistent with the possibility that analogical reasoning is involved in comprehending more creative metaphors, such as those that poets produce. But comprehension of a wide range of metaphors appears to also depend on reasoning processes that operate at the level of lexical semantics. Some form of conceptual combination may be involved (e.g., Kintsch, 2000). Rather than simply reflecting ease of retrieving vocabulary items, crystalized verbal intelligence appears to involve active integration of semantic knowledge—what I. A. Richards (1936) called "the interanimation of words".

Acknowledgments

Preparation of this paper was supported by two grants to Dušan Stamenković: Fulbright Visiting Scholar Grant (PS00232724), and a project grant from the Ministry of Education, Science, and Technological Development of the Republic of Serbia (179013); and by a Google Faculty Research Award to Keith Holyoak. We would like to thank our colleagues and research assistants for helping us in this venture: Airom Bleicher, Cristian Ramos, Kimberly Park, Jocelyn Reyes, Yasemin Yahni, Mercan Petek Kuscu, Lucia Harley, Alexa Hernandez, Jiawen Yu, and Angus Tse.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30–60.

Arthur Jr., W., Tubre, T. C., Paul, D. S., & Sanchez-Ku, M. L. (1999). College-sample psychometric and normative data on a short form of the Raven Advanced Progressive Matrices Test. *Journal of Psychoeducational Assessment*, *17*, 354–361.

Black, M. (1962). Metaphor. In M. Black, *Models and metaphors* (pp. 38–47). Ithaca, NY: Cornell University Press.

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language: A quantitative meta-analysis of neuroimaging studies of metaphor, idiom, and irony processing. *Neuropsychologica*, *50*, 2669–2683.

Bowdle, B., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*, 193–216.

Cardillo, E. R., Watson, C. E., Schmidt, G. L., Kranjec, A., & Chatterjee, A. (2012). From novel to familiar: Tuning the brain for metaphors. *NeuroImage*, *59*, 3212–3221.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. New York: Houghton Mifflin.

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

Dunbar, K. N., & Klahr, D. (2012). Scientific thinking and reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 701–718). New York: Oxford University Press.

Gagliano, A., Paul, E., Booten, K., & Hearst, M. A. (2016). Intersecting word vectors to take figurative language to new heights. In *Proceedings of the Fifth Workshop on Computational Linguistics for Literature, NAACL-HLT 2016* (pp. 20–31). San Diego, CA: Association for Computational Linguistics.

Gentner, D., & Clement, C. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. In G. H. Bower (Ed.), *Advances in the Psychology of Learning and Motivation*, Vol. 22 (pp. 307–358). New York: Academic Press.

Glucksberg, S., & Haught, C. (2006). On the relation between metaphor and simile: When comparison fails. *Mind & Language*, *21,* 360–378.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, *97*, 3–18.

Green, A. E., Kraemer, D. J., Fugelsang, A. J., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex*, *20*, 70–76.

Green, A. E., Kraemer, D. J., Fugelsang, J. A., Gray, J. R., Dunbar, K. N. (2012). Neural correlates of creativity in analogical reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 264–272.

Holyoak, K. J. (1982). An analogical framework for literary interpretation. *Poetics*, *11*, 105–126.

Holyoak, K. J. (in press). *The spider's thread: Metaphor in the mind, brain and poetry*. Cambridge, MA: MIT Press.

Holyoak, K. J., & Stamenković D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin*. http://dx.doi.org/10.1037/bul0000145

Katz, A., Paivio, A., Marschark, M., & Clark, J. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity*, *3*, 191–214.

Kazmerski, V. A., Blasko, D. G., & Dessalegn, B. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory & Cognition*, *31*, 673–689.

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*, 257–266.

Kintsch, W. (2001). Predication. *Cognitive Science*, *25*, 173–202.

Kintsch, W., & Bowles, A. R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, *17*, 249–262

Nippold, M., & M. Sullivan (1987). Verbal and perceptual analogical reasoning and proportional metaphor comprehension. *Journal of Speech and Hearing Research*, *30*, 367–376.

Pierce, R. S., & Chiappe, D. L. (2008). The roles of aptness, conventionality, and working memory in the production of metaphors and similes. *Metaphor and Symbol*, *24*, 1–19.

Rapp, A. M., Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *Neuroimage*, *62*, 600–610.

Raven, J. C. (1938). *Progressive Matrices: A perceptual test of intelligence, 1938, individual form*. London: Lewis.

Richards, I. A. (1936). *The philosophy of rhetoric.* New York: Oxford University Press.

Snow, R. E., Kyllonen, C. P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence, Vol. 2* (pp. 47–103). Hillsdale, NJ: Erlbaum.

Spearman, C. (1927). *The abilities of man*. New York: MacMillan.

Tourangeau, R., & Sternberg, R. J. (1981). Aptness in metaphor. *Cognitive Psychology*, *13*, 27–55.

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

Trick, L., & Katz, A. (1986). The domain interaction approach to metaphor processing: Relating individual differences and metaphor characteristics. *Metaphor and Symbolic Activity*, *1*, 185–213.

Vartanian, O. (2012). Dissociable neural systems for analogy and metaphor: Implications for the neuroscience of creativity. *British Journal of Psychology*, *103*, 302–316.