Metaphor and Symbol

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hmet20

Individual Differences in Comprehension of Contextualized Metaphors

Dušan Stamenković , Nicholas Ichien & Keith J. Holyoak

To cite this article: Dušan Stamenković , Nicholas Ichien & Keith J. Holyoak (2020) Individual Differences in Comprehension of Contextualized Metaphors, Metaphor and Symbol, 35:4, 285-301

To link to this article: https://doi.org/10.1080/10926488.2020.1821203

Published online: 28 Oct 2020.

METAPHOR AND SYMBOL
2020, VOL. 35, NO. 4, 285–301
https://doi.org/10.1080/10926488.2020.1821203

Individual Differences in Comprehension of Contextualized Metaphors

Dušan Stamenković[a], Nicholas Ichien[b], and Keith J. Holyoak[b]

[a]University of Niš; [b]University of California, Los Angeles

ABSTRACT
We report a study examining the role of linguistic context in modulating the influences of individual differences in fluid and crystalized intelligence on comprehension of literary metaphors. Three conditions were compared: no context, metaphor-congruent context, and literal-congruent context. Relative to the baseline no-context condition, the metaphor-congruent context facilitated comprehension of the metaphorical meaning whereas the literal-congruent context impaired it. Measures of fluid and crystalized intelligence both made separable contributions to predicting metaphor comprehension. The metaphor-congruent context selectively increased the contribution of crystalized verbal intelligence. These findings support the hypothesis that a supportive linguistic context encourages use of semantic integration in interpreting metaphors.

Introduction

A metaphor aims to illuminate a thing by describing it in terms of another thing quite unlike it. Consider this example from Vincent van Gogh (1882): "Conscience is a man's compass, and though the needle sometimes deviates, though one often perceives irregularities when directing one's course by it, one must try to follow its direction." Here the *target* concept (what is being talked about) is a conscience, and the *source* (used to characterize the target) is a compass.[1] Though a conscience is not a literal compass, the parallels between the functions of a compass as a useful if imperfect geographical guide help to make a point about the functions of a conscience as a moral guide for navigating the "path" of life. More generally, a metaphor links a target that is unfamiliar, abstract, or poorly understood to a source that is familiar, concrete, or better understood. In addition to attribution of salient features of the source domain to the target, metaphors typically result in the attribution of emergent properties that are not necessarily associated with the source and target in isolation; moreover, these properties are not always derivable by rules of semantic composition (Wilson & Carston, 2006). Metaphor is widely acknowledged to be a core manifestation of human language and thought (e.g., Lakoff & Johnson, 1980). But despite decades of research, no consensus has been reached about the psychological mechanisms that support comprehension of metaphors (for a recent review, see Holyoak & Stamenković, 2018). Although most metaphors occur within a particular linguistic or pragmatic context, relatively little work has examined the role played by context in metaphor processing. In addition, although evidence suggests that people differ in their ability to process metaphors, the nature of these individual differences remains unclear. The present paper is, to the best of our knowledge, the first to examine individual differences in the comprehension of contextualized metaphors.

CONTACT Dušan Stamenković dusan.stamenkovic@filfak.ni.ac.rs Faculty of Philosophy & Center for Cognitive Sciences, University of Niš, Niš 18000, Serbia

Note: The research data for this paper is available in Mendeley Data: https://dx.doi.org/10.17632/gykynsmmvc.1

[1]The terms *target* and *source* are commonly used in work on analogy, as well as in cognitive linguistics (Lakoff & Turner, 1989). In linguistics, the corresponding terms *tenor* and *vehicle* are often used.

Models of metaphor processing

The capacity to produce and comprehend metaphors is a distinctively human ability that depends on the interplay between language and thought. For over half a century, psychologists have attempted to develop models of metaphor processing based on some mix of general reasoning and semantic knowledge. Two broad proposals have dominated theoretical discussions. A view that traces back to Aristotle assumes that metaphor comprehension requires *analogical reasoning* to relate the target to the source (Gentner & Clement, 1988; Tourangeau & Sternberg, 1981, 1982; Trick & Katz, 1986). Black's (1962) *interaction* theory of metaphor also assumed that metaphor is based at least in part on analogy, postulating "interactions between two systems, grounded in analogies of structure (partly created, partly discovered)" (Black, 1977/1993, p. 39). Although analogies can be posed verbally, analogical reasoning can also be applied to formal, nonverbal content such as geometric patterns (e.g., Mulholland, Pellegrino, & Glaser, 1980) and mathematical problems (e.g., Novick & Holyoak, 1991). The dominant analogy-based model of metaphor (Bowdle & Gentner, 2005; Gentner & Bowdle, 2008) claims that metaphor comprehension (at least for novel metaphors) involves the application of a domain-general process for finding systematic mappings between elements of the source and target situations (e.g., for van Gogh's metaphor, emotions triggered by one's conscience may be mapped to the indicator on a compass). Given that analogical mapping appears to depend on active manipulation of complex propositional structures, and consideration of multiple alternative hypotheses, this process is likely to place a high load on working memory and executive functions such as attentional control (e.g., Cho, Holyoak, & Cannon, 2007; Waltz, Lau, Grewal, & Holyoak, 2000).

A second general approach to metaphor comprehension focuses on potential mechanisms internal to the linguistic system. A number of theorists have proposed that metaphor comprehension is based on processes of *semantic integration* continuous with those involved in literal language comprehension. In general terms, semantic integration is the incremental construction of the meaning of a message from the meanings of individual words, guided by both syntactic structure and the pragmatic context of the utterance (e.g., Bransford & Franks, 1971; Kintsch & Mangalath, 2011). Semantic integration is a broad concept, which can be viewed as continuous with the construction of situation models during text comprehension (Kintsch & van Dijk, 1978). It encompasses more specific processes such as *conceptual combination* (e.g., integrating the meanings of the words in a noun-noun compound; Wisniewski & Gentner, 1991) and *predication* (e.g., integrating the meanings of the words in a noun-verb expression; Kintsch, 2001). The well-known categorization theory of metaphor (Glucksberg & Keysar, 1990) can be viewed as a special case of predication that applies to sentences in the "is a" syntactic form. Whereas analogical reasoning is typically viewed as a domain-general process that operates on complex propositional structures held in working memory (Holyoak, 2012), semantic integration appears to be based on processes less dependent on executive functions and more dependent on semantic processing (e.g., operations on vectors of semantic features) that operate at the level of lexical semantics (Keane & Costello, 2001).

A particularly influential proposal, introduced by Glucksberg and Keysar (1990) and extended by others (e.g., Gernsbacher, Keysar, Robertson, & Werner, 2001; Glucksberg & Haught, 2006; Jones & Estes, 2005, 2006; McGlone & Manfredi, 2001), claims that metaphors are interpreted as categorization statements, in which the meaning of the source is contextually adjusted to create a novel category. For example, in van Gogh's metaphor, the literal meaning of compass as a navigational instrument might be modified to form a more abstract category of "directional guide". As noted above, metaphor-as-categorization can be modeled as a case of predication (Kintsch, 2000, 2001; Kintsch & Bowles, 2002). In Kintsch's predication model, a process of spreading activation activates concepts that link the source (expressed as the object of a linking verb, e.g., "is a man's compass") and target (e.g., "conscience"). The semantic representation of the source concept is then modified to create a new semantic representation, which in turn shades the meaning of the target.

The relative contributions of analogical reasoning and semantic integration to metaphor processing remain uncertain. It has been suggested that novel metaphors are interpreted using analogy, whereas more familiar metaphors are processed as category statements (Bowdle & Gentner, 2005). However, neural evidence suggests that metaphor processing does not primarily depend on brain regions associated with analogical reasoning. Neuroimaging studies have found that metaphor comprehension typically involves activation of areas within the temporal cortex also associated with comprehension of literal language (for meta-analyses see Bohrn, Altmann, & Jacobs, 2012; Rapp, Mutschler, & Erb, 2012; Vartanian, 2012), with additional right-hemisphere involvement in the case of novel metaphors (e.g., Arzouan, Goldstein, & Faust, 2007). By contrast, the type of complex relational integration involved in analogical reasoning is associated with activation of the left rostrolateral prefrontal cortex (for meta-analyses see Vartanian, 2012; Hobeika, Diard-Detoeuf, Garcin, Levy, & Volle, 2016; see also Bunge, Helskog, & Wendelken, 2009). This area typically is not activated during metaphor tasks (but see Prat, Mason, & Just, 2012), even for novel metaphors (Cardillo, Watson, Schmidt, Kranjec, & Chatterjee, 2012). In another dissociation, people with autism tend to show deficits in metaphor comprehension (Kalandadze, Bambina, & Næss, 2019; Morsanyi, Stamenković, & Holyoak, 2020a), even though analogical reasoning is an area in which they exhibit relative strength (Morsanyi, Stamenković, & Holyoak, 2020b).

A significant caveat on any conclusions about the processes underlying metaphor comprehension is that the novel metaphors used in research have generally been constructed by psychologists. Relative to familiar metaphors, these novel metaphors have tended to be less apt (i.e., the source describes the target less fully and accurately; Bowdle & Gentner, 2005; Jones & Estes, 2006). It is possible that literary metaphors – those created by poets and other authors – have different qualities than those constructed by psycholinguists and cognitive psychologists. A large set of literary metaphors was collected and normed by Katz, Paivio, Marschark, and Clark (1988). This item set has been extensively analyzed by Jacobs and Kinder (2017, 2018) using machine-learning algorithms and other quantitative methods. Those metaphors in the Katz et al. collection that were created by poets differ in many ways from those created by metaphor researchers. Although these differences are often subtle, machine-learning algorithms are able to distinguish literary from nonliterary metaphors with high accuracy. Jacobs and Kinder (2018) found that qualities distinguishing literary metaphors rated high in goodness include high surprisal (a statistical measure of the unexpectedness of words), relative dissimilarity of source and target concepts, the combination of concrete words with relatively complex grammar and high lexical diversity, and extra difficulty in comprehending the metaphorical meaning (see also Holyoak, 2019; Lakoff & Turner, 1989). These properties collectively suggest that good literary metaphors are high in cognitive complexity, which may be more likely to elicit analogical reasoning.

Comprehension of contextualized metaphors

A limitation of research on metaphor comprehension is that most studies have examined processing of relatively simple metaphors (typically in the nominal syntactic form, <*noun*> *is* <*noun*>) in the absence of a broader linguistic or pragmatic context. Much less is known about how metaphors are interpreted when a context is provided (Kövecses, 2015, pp. x–xi), especially in the case of literary metaphors. Note that van Gogh did not simply say, "Conscience is a man's compass;" rather, he went on to elaborate on some key correspondences. Moreover, in the passage *preceding* the metaphor, he wrote: "I think it one's duty to try to do the right thing, even knowing that one cannot go through life without making mistakes, without regret or sorrow. Somewhere I read, *Some* good *must* come by clinging to the right." The metaphor was thus introduced as the culmination of a theme already established by the prior context. Gibbs and Gerrig (1989) observed that we rarely encounter metaphors in isolation, but rather in the context of a conversation, or a piece of poetry or prose. Metaphors are generally part of some larger discourse context that provides a conceptual framework that helps a reasoner take *what is said* by a metaphorical expression and generate an interpretation of *what is meant* by that expression. Metaphor is considered to be a deeply context-sensitive communicative phenomenon (Camp, 2006).

Based on the results of several psycholinguistic studies, Carston (2012) argued that metaphorical interpretations are derived automatically when the context makes them accessible, and that people do not take longer to comprehend metaphors than literal language when metaphors are properly contextualized. Experimental evidence has indeed established that a prior context congruent with a metaphorical interpretation indeed facilitates comprehension (Gerrig & Healy, 1983; Gildea & Glucksberg, 1983; Inhoff, Lima, & Carroll, 1984; Ortony, Schallert, Reynolds, & Antos, 1978), whereas a prior context congruent with an alternative literal interpretation interferes with metaphor comprehension (McGlone & Manfredi, 2001). Shinjo and Myers (1987) found that contextual primes influence literal and metaphorical sentences in a similar fashion, and that relatively specific contexts are most effective. A recent neuroimaging study found no differences between the neural processing of metaphors and matched literal sentences when both were surrounded by a suitable context (Hartung et al., 2020). Providing a contextualized metaphorical setting can facilitate the formation of metaphor-based schemas, which are used to link elements within a text representation (Allbritton, 1995; Allbritton, McKoon, & Gerrig, 1995). Several other studies have confirmed the beneficial effect of congruent context on understanding figurative language (e.g., Gagné, Friedman, & Faries, 1996; Harris, 1976; Nayak & Gibbs, 1990; Pexman, Ferretti, & Katz, 2000; Rumelhart, 1979/1993; Thibodeau & Durgin, 2008; for a philosophical account of metaphor in context, see Stern, 2000). In general, longer, more specific, and more thematically congruent contexts provide the greatest benefit for metaphor comprehension. It is less clear *how* context exerts its effects (e.g., by facilitating either analogical reasoning or semantic integration, or by shifting the balance between these two general mechanisms). Given that a congruent linguistic context will tend to include words that share semantics with the metaphorical meaning, perhaps the most plausible hypothesis is that a congruent context facilitates semantic integration for the metaphorical meaning (Lemaire & Bianco, 2003).

Individual differences in metaphor comprehension

A relatively small number of studies have examined the impact of individual differences in cognitive abilities on metaphor comprehension (e.g., Chiappe & Chiappe, 2007; Kazmerski, Blasko, & Dessalegn, 2003; Olkoniemi, Ranta, & Kaakinen, 2016; Trick & Katz, 1986). Based on classical theories of intelligence (Cattell, 1971), two partially separable factors likely to be relevant to metaphor processing are *fluid* and *crystalized* intelligence. Fluid intelligence (closely related to individual differences in executive functions) involves reasoning (often nonverbal) about novel problems detached from prior knowledge, and is important for success in explicit analogical reasoning (Holyoak, 2012). Crystalized intelligence involves reasoning (typically verbal) that draws upon prior knowledge, and hence is likely to impact semantic integration, which depends on knowledge of lexical semantics.

Metaphor comprehension seems likely to depend on both forms of intelligence, and several approaches to metaphor comprehension based on individual differences have confirmed this general hypothesis. Chiappe and Chiappe (2007) found that individuals who scored high on a working-memory test generated higher-quality interpretations of metaphors more quickly. In addition, measures of inhibitory control (based on Stroop interference and intrusion errors on a memory test) 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 (2007) found that measures of vocabulary knowledge and exposure to printed text (linked to crystalized intelligence) also predicted metaphor quality. Indeed, crystalized intelligence yielded somewhat higher correlations with metaphor interpretation and production than did measures of working memory. Overall, these findings suggest that both fluid and crystalized intelligence have an impact on metaphor processing. In addition, studies by Silvia and Beaty (2012; Beaty & Silvia, 2013) indicate that fluid intelligence is strongly related to creativity in a metaphor production task, while crystalized knowledge predicts the ability to generate conventional metaphors.

Stamenković, Ichien, and Holyoak (2019) took an individual-differences approach to examine the comprehension of metaphors drawn from literary as well as nonliterary sources. The literary metaphors were selected from the set collected and normed by Katz et al. (1988), and extensively analyzed by Jacobs and Kinder (2017, 2018) using machine-learning algorithms. Stamenković et al. measured fluid intelligence using a short form of the nonverbal Raven's Progressive Matrices test (Arthur, Tubre, Paul, & Sanchez-Ku, 1999; Raven, 1938). Crystalized verbal intelligence was assessed using a novel Semantic Similarities Test (SST), in which word pairs are presented with the question, "How are the two concepts in each pair similar to one another?" Scores on the SST correlated well with the Vocabulary subscale of the Wechsler Adult Intelligence Scale, a standard measure of crystalized intelligence. For nonliterary metaphors, regression analyses indicated that only the measures of crystalized intelligence predicted unique variance in comprehension scores. In contrast, for literary metaphors measures of both fluid and crystalized intelligence made separable contributions to predicting comprehension. This pattern suggests that analogical reasoning (known to be strongly linked to fluid intelligence) may play an important role in comprehension only for relatively complex metaphors, whereas some form of semantic integration (more dependent on crystalized intelligence) influences metaphor comprehension across a broader range of metaphor types (nonliterary as well as literary).

Overview of present study

The present study sought to determine whether and how individual differences in fluid and crystalized intelligence impact the comprehension of literary metaphors when a prior linguistic context is provided. We sought to extend the approach of Stamenković et al. (2019), who only examined metaphors presented as isolated sentences, without a context. We focused on literary metaphors because our earlier findings indicated that for these cognitively complex metaphors, in the absence of a context both types of cognitive abilities contribute to prediction of comprehension.

Here we assessed comprehension of a metaphor (presented as a critical sentence) in three conditions: (1) *no context* (a baseline condition comparable to that tested by Stamenković et al., 2019), (2) *metaphor-congruent context* (preceding text supporting the metaphorical interpretation of the critical sentence), and (3) *literal-congruent* (preceding text focusing on target only, rather than supporting the metaphorical interpretation of the critical sentence). The third condition is similar to what previous researchers (Keysar, Shen, Glucksberg, & Horton, 2000; Thibodeau & Durgin, 2008) have termed a "no-mapping" or "non-metaphorical" scenario.

Based on previous work on context effects we predicted that, relative to the baseline no-context condition, a metaphor-congruent context would facilitate comprehension of the metaphorical meaning whereas a literal-congruent context would impair it. Our central aim was to assess how the context variations would influence the pattern of individual differences. Based on the previous findings of Stamenković et al. (2019) for literary metaphors, we expected measures of both fluid and crystalized intelligence to make separable contributions to predicting metaphor comprehension, at least in the no-context condition. We hypothesized that the metaphor-congruent context would encourage and facilitate semantic integration as the dominant strategy for interpreting the critical sentence, perhaps resulting in greater reliance on crystalized intelligence. It was less clear how the literal-congruent context would impact individual differences.

Method

Participants

Participants were 201 undergraduates at the University of California, Los Angeles (UCLA) (female = 152, male = 48, genderqueer = 1; mean age = 19.3) who received course credit. The study was approved, including informed consent procedures, by the UCLA Office of the Human Research Protection Program. This sample size is consistent with those used in similar studies focusing on individual differences in metaphor comprehension (Stamenković et al., 2019), relational similarity (Ichien, Lu, & Holyoak, 2019), and relational reasoning (Gray & Holyoak, 2020). The participants were mainly native speakers of English (79%), with a minority of bilinguals who spoke English fluently (21%, self-assessed). We adopted the exclusion criteria based on those used by Stamenković et al. (2019) to identify carelessness or inattention on verbal tasks. Following these criteria, we did not analyze data from an additional 15 participants who either scored below 12 (max = 40) on the Semantic Similarities Test, below 10 (max = 40) on the metaphor test, or failed to complete the entire battery of tests. After these exclusions, a total of 65 participants were assigned to the no-context condition, 69 to the metaphor-congruent condition, and 67 to the literal-congruent condition.

Design, materials, and procedure

All participants completed four tasks in a fixed order. The first three were assessments of individual differences on cognitive tasks, and the fourth and final task involved metaphor comprehension. The task order was identical to that used by Stamenković et al. (2019, Study 3); the order was constant so as to minimize variability in the predictive tasks (as our focus was on individual differences). All tasks were administered to participants individually by computer. None of the tasks involved any sort of time pressure and participants could take breaks between tasks. The entire test session lasted approximately 50 minutes to an hour.

Task 1: *Raven's Progressive Matrices (RPM)*. A short version of the RPM (Arthur et al., 1999), adapted for computer administration using SuperLab software, was administered. Each matrix problem consists of a 3 × 3 array in which geometric forms vary systematically across rows and columns. The lower right cell is empty, and the task is to select the best completion from a set of eight options. The RPM is thus similar to geometric analogy problems, and is generally considered a central measure of fluid intelligence (Snow, Kyllonen, & Marshalek, 1984). In addition to predicting performance on psychometric analogy tests, scores on the RPM correlate with production of relational responses in a picture-mapping task (Vendetti, Wu, & Holyoak, 2014), spontaneous analogical transfer in a problem-solving task (Kubricht, Lu, & Holyoak, 2017), and similar tasks requiring relational reasoning (Gray & Holyoak, 2020). To the extent that explicit analogical reasoning is required to comprehend metaphors, the RPM should be a robust predictor of individual differences in metaphor comprehension.

Task 2: *Semantic Similarities Test (SST)*. This test, introduced by Stamenković et al. (2019), provides 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 pairs of single words, where the similarities vary in degree of abstraction. The test was designed to be similar to the Wechsler Adult Intelligence Scale (WAIS) Similarities subscale (a measure of verbal comprehension), but uses entirely different items, selected to span a broader range of similarities. The SST is a verbal test in which semantic knowledge of word meanings is critical. In addition to predicting comprehension of both literary and nonliterary metaphors (Stamenković et al., 2019), the SST has been found to predict performance on a battery of relational reasoning tasks (Gray & Holyoak, 2020).

Task 3: *Vocabulary Subtest of Wechsler Adult Intelligence Scale*. The third test was the Vocabulary subtest of the third version of the Wechsler Adult Intelligence Scale (WAIS-III), a standardized measure of crystalized intelligence consisting of 33 vocabulary items. Responses were assessed using the official scoring manual (Wechsler, 1997).

Task 4: *Literary metaphor comprehension*. The final task consisted of 20 literary metaphor comprehension items selected from a list of literary metaphors drawn from poetry anthologies by Katz et al. (1988), and rated high on their "goodness" scale. Each metaphor was in the form of a single critical sentence (e.g., *Nerves after a quarrel are frozen leaves in winter*; see Table 1). Their syntactic forms included nominal (<*noun*> *is* <*noun*>), nominal with an adjective modifier, and nominal with

Table 1. Context conditions and scoring key for an example of a literary metaphor: *Nerves after a quarrel are frozen leaves in winter.*

Context Text		
	Metaphor-congruent Condition	**Literal-congruent condition**
Preceding text	We were fighting from early in the morning until lunch. Many awful things were said, and at some points I could not believe what was happening. Now that it has finally stopped, my whole body is shaking. I can't really think reasonably.	A quarrel is an angry argument or disagreement. There are people who fight all the time, and it's sometimes hard to understand why. It's simply part of who they are and what they are. There is no point in trying to understand such people.
Scoring Key		
Points	**2 points (completely correct)**	**1 point (partially correct)**
Sample responses	After a quarrel, it is difficult to do anything. A person may be deeply affected by arguments and confrontation. Nerves are preventing one from fully thinking about the situation.	Nerves leave you speechless after fighting. After arguing a person is shocked. Nerves after a fight can be stunning and make you feel paralyzed/numb.

a prepositional phrase (for detailed analyses of the properties of this metaphor set, see Jacobs & Kinder, 2018). All the literary metaphors used in the study are provided in Supplemental Online Materials, Appendix I.

The metaphors were presented one at a time on a computer display, in a different random order for each participant. Comprehension was assessed by an open-ended task requiring participants to type in their interpretation of each of the metaphorical statements. An initial example was provided. In the no-context condition, participants were simply instructed to "tell us how you understand the meaning" of each expression, responding by typing their answer in a box on the computer screen. For the two context conditions, the full passage (context followed by the critical sentence, which was underlined) was shown on the screen. Participants were instructed, "Your task is to read each text and the ending expression and tell us how you understand the meaning of the underlined sentence." Responses were entered in the same way as for the no-context condition.

Participants were randomly assigned to one of three between-subjects conditions for the metaphor task (see Table 1; complete context materials are provided in Supplemental Online Materials, Appendix II). In the *no-context* condition all metaphors were presented without any context. In the *metaphor-congruent* condition, for each critical sentence a prior linguistic context established a theme supporting the intended metaphorical interpretation. These contexts were composed by the authors. In the *literal-congruent* condition, the linguistic context was designed to support a literal interpretation of the critical sentence, and thus to interfere with finding the metaphorical interpretation. *Literal-congruent* contexts provided encyclopedia-like information related to topics relevant to the target domain.

Results

Metaphor comprehension scores

Responses to the metaphor task were graded using a scoring key (see Table 1; also Supplemental Online Materials, Appendix I). The scoring key was constructed by two raters (D.S. and N.I.) based on a sample of 127 initial participant responses. A response was awarded 0, 1, or 2 points. A score of 2 (completely correct) was given if the paraphrase captured the metaphorical meaning at a level of abstraction beyond the source domain (i.e., a paraphrase that did not repeat the metaphorical formulation). A score of 1 (partially correct) was given if the paraphrase described the metaphorical meaning but maintained explicit links to the source domain. A score of 0 (incorrect) was given if the paraphrase was strongly linked to the source domain, merely ascribed simple physical features to the target, or rephrased the metaphor using simpler terms with no further insights (i.e., was literal in nature, constituting 32.1% of paraphrases receiving scores of 0), was a complete misinterpretation or nonsensical (62.6% of 0 scores), or if no response was entered (5.3% of 0 scores). The data for all three conditions were collected simultaneously; all raters graded responses as a single undifferentiated batch, and were thus blind to condition, as well as to the results of the other tasks. Examples of different types of errors are provided in Supplemental Online Materials, Appendix III. Inter-rater agreement for the sample of responses scored by two raters was 95%. After the scoring key had been established, all other responses were scored by D.S., and all scoring was blind to condition. We calculated the mean score per item, thus placing the means on the 0–2 range used in scoring each individual item. These mean comprehension scores were 1.03 ($SD = .21$, skewness = .30, kurtosis = .59) for the no-context baseline condition, 1.18 ($SD = .20$, skewness = .29, kurtosis = .58) for the metaphor-congruent context, and 0.86 ($SD = .26$, skewness = .29, kurtosis = .57) for the literal-congruent context. Shapiro-Wilks tests of normality for comprehension scores within each condition revealed no violation of normality for any condition (no-context baseline: $W(65) = .98$, $p = .323$; metaphor-congruent context: $W(69) = .97$, $p = .109$; literal-congruent context: $W(67) = .98$, $p = .227$).

Individual difference measures

Table 2 summarizes the descriptive statistics for the three measures we obtained (RPM, SST, and WAIS-III Vocabulary), and Table 3 shows the intercorrelations among them. As would be expected given random assignment of participants, the pattern did not differ reliably across context conditions for any of these measures; hence Tables 2 and 3 present data from the entire set of 201 participants in the study. We used the *cocor* package in R to test the difference between these intercorrelations (Diedenhofen & Musch, 2015). The *cocor* package implements ten distinct statistical methods to test the difference between correlation coefficients; for our analyses, all methods produced convergent results. For convenience, we report the results implementing Dunn and Clark's (1969) z-test. The correlation between performance on our two measures of crystallized verbal intelligence, the WAIS-III Vocabulary measure and the SST ($r = .471$), was reliably greater than that between the WAIS-III and the RPM, our measure of fluid intelligence ($r = .285$, $z = 2.59$, $p = .01$), but not reliably greater than that between the SST and the RPM ($r = .367$, $z = 1.41$, $p = .16$). This pattern (also observed by Stamenković et al., 2019) is consistent with the interpretation that the RPM measures fluid intelligence whereas both the SST and WAIS-III Vocabulary tests measure crystalized intelligence. These two forms of cognitive ability are related to one another but separable (Cattell, 1971).

Influence of context and individual differences on metaphor comprehension

The next step was to examine the patterns of intercorrelations linking individual differences to metaphor comprehension across the three context conditions. For these analyses we standardized scores on all three individual-difference tests, and used the mean of the SST and WAIS-III Vocabulary scores to form a composite measure of crystalized intelligence (following the procedure employed by

Table 2. Descriptive statistics for each individual-difference measures (across all context conditions).

Test	Mean	Max	SD	Range
RPM	6.78	12	2.87	0–12
SST	28.32	37	4.00	15–37
WAIS-III Vocabulary	47.16	62	8.31	6–62

Table 3. Pearson's r correlations among individual-difference measures (across all context conditions).

Test	WAIS-III Vocabulary	SST
RPM	.285	.367
SST	.471	

All p's < .001

Stamenković et al., 2019). This combined measure allows for a more comprehensive measure of crystalized verbal knowledge than what is individually captured by the SST and WAIS-III scores. The RPM was used as a measure of fluid intelligence.

We performed a multiple regression analysis to examine the influence of both context and individual differences in fluid and crystalized verbal intelligence on metaphor comprehension within a unified model:

$$M = C_M + C_L + I_F + I_C + C_M * I_F + C_M * I_C + C_L * I_F + C_L * I_C.$$

Here, M represents metaphor comprehension score; C_M represents metaphor-congruent context, compared to the no-context baseline; C_L represents the literal-congruent context, compared to the no-context baseline; I_F represents fluid intelligence (standardized RPM scores); and I_C represents crystalized verbal intelligence (standardized mean of standardized SST and WAIS-III scores). Product terms in the equation represent interactions between factors (e.g., $C_M * I_F$ represents an interaction between metaphor-congruent context, compared to the no-context baseline, and fluid intelligence). A Levene's test on metaphor comprehension scores across conditions found no violation of homogeneity of variances, $F(2,198) = 1.65$, $p = .194$, and our unified model significantly accounted for metaphor comprehension, $F(8, 192) = 26.12$, $p < .001$, adjusted R-squared = .50.

Examining specific effects within the model, we first considered the effect of context on metaphor comprehension. The analysis revealed a main effect for metaphor-congruent context compared to the no-context baseline, $t(192) = 5.91$, $p < .001$, $\beta = .73$, as well as a main effect for literal-congruent context compared to the no-context baseline, $t(192) = 5.46$, $p < .001$, $\beta = -.68$. These results support the prediction that relative to the no-context baseline, metaphor comprehension was facilitated by a metaphor-congruent context but impaired by a literal-congruent context.

Next, we examined the effects of individual differences in fluid and crystalized verbal intelligence on metaphor comprehension scores. Figure 1 shows our individual difference measures plotted against metaphor comprehension scores for each context. There was a reliable effect of fluid intelligence on metaphor comprehension score in the no-context baseline, $t(192) = 2.04$, $p = .043$, $\beta = .20$. Further, there was no interaction between fluid intelligence and the metaphor-congruent context compared to the no-context baseline, $t(192) = 1.82$, $p = .070$, $\beta = -.24$, or for literal-congruent context compared to the no-context baseline, $t(192) = 1.11$, $p = .269$, $\beta = -.15$. These results indicate that fluid intelligence was a reliable predictor of metaphor comprehension scores, and that individual contexts did not reliably differ in the extent to which fluid intelligence predicted metaphor comprehension.

In addition, there was a reliable effect of crystalized verbal intelligence on metaphor comprehension in the no-context baseline condition, $t(192) = 2.89$, $p = .004$, $\beta = .29$. Further, there was a reliable interaction of crystalized intelligence with metaphor-congruent context compared to the no-context baseline, $t(192) = 2.66$, $p = .008$, $\beta = .37$. However, there was no reliable interaction with the literal-congruent context compared to the no-context baseline, $t(192) = 1.20$, $p = .233$, $\beta = .16$. These findings highlight the association between crystalized verbal intelligence and metaphor comprehension across contexts, and indicate that the metaphor-congruent context increased the contribution of crystalized verbal intelligence. This increased reliance on semantic knowledge suggests that the metaphor-congruent context encouraged and facilitated semantic integration as the dominant strategy for interpreting the critical sentence.

Although both individual difference measures reliably accounted for separable variance in metaphor comprehension scores across contexts, we went on to directly compare the effects of fluid intelligence and crystalized intelligence on metaphor comprehension. A Wald test revealed that crystalized verbal intelligence was a stronger predictor than fluid intelligence in the no-context baseline, $F(2,192) = 10.93$, $p < .001$. Further Wald tests revealed the same pattern of effects for the metaphor-congruent, $F(5, 192) = 19.24$, $p < .001$, and the literal-congruent contexts, $F(5, 189) = 16.61$, $p < .001$. These findings replicate those of Stamenković et al. (2019) for the no-context baseline, and extend them to metaphor-congruent and literal-congruent contexts.

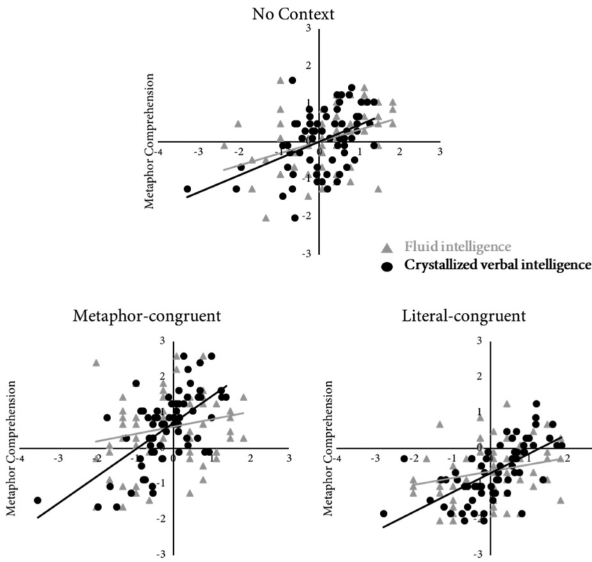

Figure 1. Individual differences in fluid intelligence (standardized RPM scores) and crystalized verbal intelligence (mean of standardized SST and WAIS-III scores) plotted against standardized metaphor comprehension scores for each context condition.

Influence of context on response types for metaphor comprehension

The fact that a literal-congruent context led to lower metaphor comprehension scores, despite the absence of any time pressure, suggests that participants were unable to apply an effective strategy to overcome the literal or misleading orientation evoked by this type of context. To provide a more detailed picture of the influence of the alternative contexts, Figure 2 shows the proportion of responses for each context condition that received scores of 0 (incorrect), 1 (partially correct), and 2 (completely correct). We conducted three one-way ANOVAs across context conditions (performed separately for 0, 1 and 2 scores). In order to control FWER at .05 across these three separate ANOVA's, we used a more stringent, adjusted alpha of .017 (.05/3 = .017). These tests yielded a significant main effect for the number of 0 scores, $F(2,198) = 31.27$, $p < .001$, $\eta^2 = 0.24$, as well as for the number of 2 scores, $F(2,198) = 22.36$, $p < .001$, $\eta^2 = 0.18$. The number of 1 scores also varied across conditions, although the effect size was smaller, $F(2,198) = 4.26$, $p = .015$, $\eta^2 = 0.04$. Relative to the no-context baseline, the metaphor-congruent and literal-congruent contexts generated mirror-image response patterns for 0 and 2 scores. Among the three conditions, the metaphor-congruent context yielded the lowest proportion of 0 scores and the highest proportion of 2 scores, whereas the literal-congruent context yielded the *highest* proportion of 0 scores and the *lowest* proportion of 2 scores.

A more detailed analysis was performed on the frequencies of the two major subtypes responses receiving a score of 0 on the metaphor comprehension task, *literal responses* and *misinterpretations*. A literal response interpreted the target metaphor as a literal sentence, where the words or phrases representing the target and/or source concepts retain their respective conventional meanings, as opposed to being combined to form an abstracted, metaphorical meaning. For example, for the target metaphor "Nerves after a quarrel are frozen leaves in winter," where "nerves after a quarrel" represents the target concept and "frozen leaves in winter" represents the source concept, literal responses included, "Frozen leaves are still but very delicate," "Nerves after fighting are cold," and "Nerves are stuck." In contrast, a misinterpretation of a target metaphor significantly diverged from both a metaphorical and a literal interpretation. For example, given the same target metaphor, misinterpretations included, "Nerves usually tend to calm down after a fight," "All your nerves just stop working when encounter something bad," and "Nerves after a quarrel are not easily changed." Further examples are provided in Supplemental Online Materials, Appendix III.

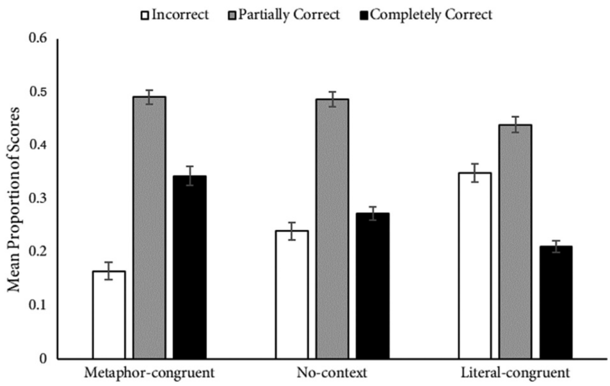

Figure 2. Proportion of responses receiving each metaphor comprehension score (0, 1 or 2) for each of the three context conditions. Error bars indicate ± standard error of the mean.

Shapiro-Wilks tests of normality for literal responses within each condition revealed violations of normality for the no-context baseline, $W(65) = .84$, $p > .001$; the metaphor-congruent context, $W(69) = .73$, $p > .001$; and the literal-congruent context, $W(67) = .95$, $p = .005$. A Kruskal-Wallace test of context on the frequency of literal responses revealed a significant difference across the three conditions, $H(2) = 70.91$, $p < .001$. Post-hoc pairwise comparisons with a Bonferroni correction to control FWER at .05 revealed that participants produced more literal responses in the literal-congruent condition (mean rank = 144.24) than in either the no-context (mean rank = 96.88, Standardized U = 4.81, $p > .001$) or metaphor-congruent conditions (mean rank = 62.89, Standardized U = 8.39, $p > .001$), and produced more literal responses in the no-context condition than in the metaphor-congruent condition (Standardized U = 3.48, $p = .002$).

Shapiro-Wilks tests of normality for misinterpretations within each condition revealed violations of normality for the no-context baseline, $W(65) = .93$, $p > .001$; the metaphor-congruent context, $W(69) = .87$, $p > .001$; and the literal-congruent context, $W(67) = .96$, $p = .043$. A Kruskal-Wallace test of the effect of context on the frequency of misinterpretations revealed a significant difference across the three conditions, $H(2) = 18.73$, $p < .001$. Post-hoc pairwise comparisons with a Bonferroni correction to control FWER at .05 revealed that participants produced more misinterpretations in the literal-congruent condition (mean rank = 123.60) than in either the no-context (mean rank = 98.94, Standardized U = 2.46, $p = .042$) or the metaphor-congruent conditions (mean rank = 80.99, Standardized U = 4.31, $p > .001$). The number of misinterpretations did not differ reliably across the no-context condition and the metaphor-congruent condition (Standardized U = 1.80, $p = .214$).

Thus, the literal-congruent condition evoked a greater frequency of both literal interpretations of metaphors and complete misinterpretations of them. This overall pattern is consistent with the hypothesis that both types of contexts encouraged a semantic integration strategy. However, whereas a metaphor-congruent context favors semantic integration consistent with the metaphorical interpretation, a literal-congruent context impairs the integration process. Sometimes this interference results in a literal interpretation, but often the inherent conflict between literal and metaphorical interpretations results in a complete misinterpretation.

Discussion

Summary and implications

The present findings reveal that the impact of individual differences on metaphor comprehension is modulated by prior context. Overall, comprehension of literary metaphors is positively related to measures of both fluid and crystalized intelligence, confirming and extending the previous findings of

Stamenković et al. (2019). When a literary metaphor is preceded by a metaphor-congruent context, the predictive power of crystalized intelligence is enhanced. These findings support the hypothesis that a supportive linguistic context encourages use of semantic integration in interpreting metaphors.

The observed patterns of individual differences have implications for understanding the cognitive mechanisms that may underlie metaphor comprehension. For several decades, two general theoretical proposals – analogical reasoning and semantic integration – have dominated discussion. Analogical reasoning appears to place substantial demands on working memory and inhibitory control – i.e., domain-general executive functions closely linked to fluid intelligence. In contrast, semantic integration is viewed as more language-specific, and particularly dependent on semantic knowledge. If integration can be performed by operations on vectors of semantic features (e.g., Kintsch & Mangalath, 2011), without necessitating explicit manipulation of relations in working memory, then the burden on executive functions will be reduced relative to the requirements for analogical reasoning.

Thus, we would expect individual differences in fluid intelligence to have a strong association with metaphor comprehension when an analogy strategy plays a prominent role, whereas individual differences in crystalized intelligence are more likely to dominate when semantic integration is the primary strategy. From previous work using metaphors presented in isolation (Stamenković et al., 2019), it appears that crystalized intelligence predicts successful comprehension for both literary and nonliterary metaphors, whereas fluid intelligence is primarily predictive for more complex literary metaphors of the sort used in the present study. These findings suggest that the role of analogical reasoning may be limited to comprehension of more complex metaphors, whereas semantic integration plays a more general role for all metaphor types. The present findings demonstrate that for literary metaphors, the impact of variations in crystalized intelligence is enhanced when the metaphor is preceded by a supportive linguistic context. Taken as a whole, the evidence from individual-difference data implies that semantic integration may be involved in grasping all types of metaphors, whether encountered in isolation or in a meaningful context, but is especially important when a supportive linguistic context is provided.

An intriguing aspect of the present findings is that even an unfavorable context – one that is congruent with a literal rather than metaphorical interpretation – apparently involves a comprehension strategy based at least in part on semantic integration. The general tendency to rely on semantic integration is consistent with the view that integration processes are fundamental to comprehension of meaningful text, whether its content is literal, metaphorical, or some mix of both (Kintsch, 1998; Kintsch & Mangalath, 2011). In other words, people naturally apply semantic integration at all times when comprehending text, supplementing this strategy with more explicit reasoning processes mainly under special circumstances (e.g., upon encountering a complex metaphor devoid of context). Our findings are consistent with those of a recent neuroimaging study that did not find differences between the neural processing of metaphors and matched literal sentences when both were embedded in an appropriate narrative context (Hartung et al., 2020).

Directions for future research

The current study has several limitations. One important issue stems from the nature of the metaphor comprehension task used in the present study. To provide correct answers on this type of open-ended metaphor comprehension task, it is necessary to both comprehend the metaphor and express its meaning verbally. It is possible that those with higher crystalized intelligence scores are better able to construct paraphrases of metaphor meaning. It would be useful to perform follow-up studies using other measures of comprehension (e.g., multiple-choice format), so as to more clearly separate comprehension ability from verbal expression.

It would also be valuable to compare the impact of different types of contexts, especially for literary metaphors. A recent study of the time course and neurophysiological underpinnings of metaphors in literary contexts suggests that such contexts enable the manipulation of multiple meanings in working memory (Bambini, Canal, Resta, & Grimaldi, 2019). Especially for novel metaphors, the simultaneous evocation of multiple meanings creates what have been termed *plurisigns* (Wheelwright, 1968), which yield poetic effects (Holyoak, 2019). A literary context may also serve to establish a literary mode that facilitates the recognition of metaphors (Steen, 1994).

The individual-differences approach has proved useful in identifying mechanisms involved in understanding literary and nonliterary metaphors; however, many important questions remain open. Literary metaphors, which reflect the most creative aspects of language, can be further differentiated based on a variety of cognitive and emotional characteristics (Holyoak, 2019; Jacobs, 2015; Jacobs & Kinder, 2018). For all types of metaphors, variations in such factors as aptness, familiarity, syntactic complexity, and source-target similarity appear to influence comprehension (see Holyoak & Stamenković, 2018; Roncero & de Almeida, 2015). Future research should attempt to tease apart the contributions of different dimensions of variation among metaphors. Such differences include variations in the syntactic forms of metaphors (e.g., Cardillo et al., 2012; Li, Zhu, & Wang, 2013; Rai & Chakraverty, 2020), as well as varying levels of abstraction (Banaruee, Khoshsima, Zare-Behtash, & Yarahmadzai, 2019).

Another valuable extension of the individual-differences approach would be to examine metaphor comprehension and production tasks with the same participants. Given the strong association that exists between fluid intelligence and working memory capacity (e.g., Kane, Hambrick, & Conway, 2005), and the role that executive functions play in metaphor comprehension (e.g., Chiappe & Chiappe, 2007; Pierce & Chiappe, 2008), future work on individual differences should include tests of working memory. The individual-differences approach can also be combined with measures of creative potential (e.g., Faust & Kenett, 2014; Kenett, Gold, & Faust, 2018) to provide further insights into literary metaphor comprehension. Other individual differences measures, such as measures of personality traits, preference for imagery, and emotional understanding (see Fetterman, Bair, Werth, Landkammer, & Robinson, 2016) could provide more comprehensive understanding of the range of factors that impact metaphor comprehension and production. It would also be extremely interesting to compare the performance of different age groups. Although fluid intelligence is reduced with normal aging, the type of semantic knowledge underlying crystalized intelligence is spared or even enhanced (Salthouse, 2004). Thus older adults may maintain their abilities in metaphor processing, even though analogical reasoning with formal materials is impaired (Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004).

In addition, future research should consider metaphorical processing in languages other than English. Attention to cross-linguistic differences may prove particularly revealing, given the emerging picture of metaphor comprehension as heavily dependent on language-specific processes encompassed by semantic integration. Cognitively relevant differences across particular languages in their syntax or lexicon may offer useful avenues to illuminate how human reasoners combine lexical units to produce metaphorical speech or text, and to extract contextualized meaning from such speech or text. Another direction of future research within the individual-differences approach could relate it to pictorial and multimodal metaphors.

A deeper understanding of the roles played by the many factors that distinguish metaphors within a language as well as across languages will require the development of new computational models of metaphor processing (for a recent survey of computational approaches to metaphor comprehension, see Rai & Chakraverty, 2020). In this regard, recent work in artificial intelligence may prove to be relevant. In particular, machine-learning algorithms applied to large text corpora are able to create feature vectors representing the meanings of individual words (e.g., Mikolov, Sutskever, Chen, Corrado, & Dean, 2013), as well as of larger linguistic units such as sentences (Devlin, Chang, Lee, & Toutanova, 2018). These tools, which generate feature-based semantic representations that can be merged to create contextualized meanings, may prove to be applicable to some aspects of semantic integration involved in metaphor comprehension.

Acknowledgments

Preparation of this paper was supported by the Faculty of Philosophy (University of Niš) internal grant No 360/1-16-1-01 to Dušan Stamenković and by NSF Grant BCS-1827374 to Keith Holyoak. We thank Aliza Ayaz, Diego Fernandez-Salvador, and Anshu Patel for assistance in data collection.

Disclosure statement

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding

This work was supported by the NSF [BCS-1827374]; Faculty of Philosophy (University of Niš) internal grant [360/1-16-1-01].

ORCID

Dušan Stamenković http://orcid.org/0000-0002-0121-4591
Nicholas Ichien http://orcid.org/0000-0002-0928-0809
Keith J. Holyoak http://orcid.org/0000-0001-8010-6267

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. doi:10.1037/0033-2909.131.1.30

Allbritton, D. W. (1995). When metaphors function as schemas: Some cognitive effects of conceptual metaphors. *Metaphor and Symbolic Activity*, *10*, 33–46. doi:10.1207/s15327868ms1001_4

Allbritton, D. W., McKoon, G., & Gerrig, R. J. (1995). Metaphor-based schemas and text representations: Making connections through conceptual metaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 612–625. doi:10.1037/0278-7393.21.3.612

Arthur, W., Jr., 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. doi:10.1177/073428299901700405

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, *1160*, 69–81. doi:10.1016/j.brainres.2007.05.034

Bambini, V., Canal, P., Resta, D., & Grimaldi, M. (2019). Time course and neurophysiological underpinnings of metaphor in literary context. *Discourse Processes*, *56*, 77–97. doi:10.1080/0163853X.2017.1401876

Banaruee, H., Khoshsima, H., Zare-Behtash, E., & Yarahmadzai, N. (2019). Types of metaphors and mechanisms of comprehension. *Cogent Education*, *6*(1), article 1617824. doi:10.1080/2331186X.2019.1617824

Beaty, R. E., & Silvia, P. J. (2013). Metaphorically speaking: Cognitive abilities and the production of figurative language. *Memory & Cognition*, *41*(2), 255–267. doi:10.3758/s13421-012-0258-5

Black, M. (1962). Metaphor. In M. Black (Ed.), *Models and metaphors* (pp. 25–47). Ithaca, NY: Cornell University Press.

Black, M. (1977/1993). More about metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.) (pp. 19–43). New York: Cambridge 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. doi:10.1016/j.neuropsychologia.2012.07.021

Bowdle, B., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*, 193–216. doi:10.1037/0033-295X.112.1.193

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *2*, 331–350. doi:10.1016/0010-0285(71)90019-3

Bunge, S. A., Helskog, E. H., & Wendelken, C. (2009). Left, but not right, rostrolateral prefrontal cortex meets a stringent test of the relational integration hypothesis. *NeuroImage*, *46*, 338–342. doi:10.1016/j.neuroimage.2009.01.064

Camp, E. (2006). Contextualism, metaphor, and what is said. *Mind & Language*, *21*, 280–309. doi:10.1111/j.1468-0017.2006.00279.x

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. doi:10.1016/j.neuroimage.2011.11.079

Carston, R. (2012). Metaphor and the literal/non-literal distinction. In K. Allan & K. M. Jaszczolt (Eds.), *The Cambridge handbook of pragmatics* (pp. 469–492). Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. New York, NY: 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. doi:10.1016/j.jml.2006.11.006

Cho, S., Holyoak, K. J., & Cannon, T. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition*, *35*, 1445–1455. doi:10.3758/BF03193614

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of deep bidirectional transformers for language understanding*. Retrieved from http://arxiv.org/abs/1810.04805

Diedenhofen, B., & Musch, J. (2015). Coco: A comprehensive solution for the statistical comparison of correlations. *PLoS One*, *10*(4), e0121945. doi:10.1371/journal.pone.0121945

Dunn, O. J., & Clark, V. A. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, *64*, 366–377. doi:10.1080/01621459.1969.10500981

Faust, M., & Kenett, Y. N. (2014). Rigidity, chaos and integration: Hemispheric interaction and individual differences in metaphor comprehension. *Frontiers in Human Neuroscience*, *8*, 511. doi:10.3389/fnhum.2014.00511

Fetterman, A. K., Bair, J. L., Werth, M., Landkammer, F., & Robinson, M. D. (2016). The scope and consequences of metaphoric thinking: Using individual differences in metaphor usage to understand how metaphor functions. *Journal of Personality and Social Psychology*, *110*(3), 458–476. doi:10.1037/pspp0000067

Gagné, C. L., Friedman, A., & Faries, J. (1996). Effect of priming on the comprehension of predicative metaphors. *Metaphor and Symbolic Activity*, *11*, 125–143. doi:10.1207/s15327868ms1102_2

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, NY: Academic Press.

Gentner, D., & Bowdle, B. (2008). Metaphor as structure-mapping. In R. W. Gibbs Jr. (Ed.), *Cambridge handbook of metaphor and thought* (pp. 109–128). New York: Cambridge University Press.

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. doi:10.1006/jmla.2000.2782

Gerrig, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 667–675. doi:10.1037/0278-7393.9.4.667

Gibbs, R. W., Jr., & Gerrig, R. J. (1989). How context makes metaphor comprehension seem 'special'. *Metaphor and Symbolic Activity*, *4*, 145–158. doi:10.1207/s15327868ms0403_3

Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior*, *22*, 577–590. doi:10.1016/S0022-5371(83)90355-9

Glucksberg, S., & Haught, B. (2006). On the relation between metaphor and simile: When comparison fails. *Mind & Language*, *21*, 360–378. doi:10.1111/j.1468-0017.2006.00282.x

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, *97*, 3–18. doi:10.1037/0033-295X.97.1.3

Gray, M. E., & Holyoak, K. J. (2020). Individual differences in relational reasoning. *Memory & Cognition*, *48*, 96–110. doi:10.3758/s13421-019-00964-y

Harris, R. J. (1976). Comprehension of metaphors: A test of the two-stage processing model. *Bulletin of the Psychonomic Society*, *8*, 312–314. doi:10.3758/BF03335150.

Hartung, F., Kenett, Y. N., Cardillo, E. R., Humphries, S., Klooster, N., & Chatterjee, A. (2020). Context matters: Novel metaphors in supportive and non-supportive contexts. *NeuroImage*, *212*, 116645. doi:10.1016/j.neuroimage.2020.116645

Hobeika, L., Diard-Detoeuf, C., Garcin, B., Levy, R., & Volle, E. (2016). General and specialized brain correlates for analogical reasoning: A meta-analysis of functional imaging studies. *Human Brain Mapping*, *37*, 1953–1969. doi:10.1002/hbm.23149

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 234–259). New York: Oxford University Press.

Holyoak, K. J. (2019). *The spider's thread: Metaphor in mind, brain, and poetry*. Cambridge, MA: MIT Press. doi:10.7551/mitpress/11119.001.0001

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin*, *144*, 641–671. doi:10.1037/bul0000145

Ichien, N., Lu, H., & Holyoak, K. J. (2019). Individual differences in judging similarity between semantic relations. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual meeting of the cognitive science society* (pp. 464–470). Austin, TX: Cognitive Science Society.

Inhoff, A. W., Lima, S. D., & Carroll, P. J. (1984). Contextual effects on metaphor comprehension in reading. *Memory & Cognition*, *12*, 558–567. doi:10.3758/BF03213344

Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers Human Neuroscience*, *9*, 186. doi:10.3389/fnhum.2015.00186

Jacobs, A. M., & Kinder, A. (2017). The brain is the prisoner of thought: A machine-learning assisted quantitative narrative analysis of literary metaphors for use in Neurocognitive Poetics. *Metaphor and Symbol*, *32*, 139–160. doi:10.1080/10926488.2017.1338015

Jacobs, A. M., & Kinder, A. (2018). What makes a metaphor literary? Answers from two computational studies. *Metaphor and Symbol*, *33*, 85–100. doi:10.1080/10926488.2018.1434943

Jones, L. L., & Estes, Z. (2005). Metaphor comprehension as attributive categorization. *Journal of Memory and Language*, *53*, 110–124. doi:10.1016/j.jml.2005.01.016

Jones, L. L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, *55*, 18–32. doi:10.1016/j.jml.2006.02.004

Kalandadze, T., Bambina, V., & Næss, K.-A. B. (2019). A systematic review and meta-analysis of studies on metaphor comprehension in individuals with autism spectrum disorder: Do task properties matter? *Applied Psycholinguistics*, *40*, 1421–1454. doi:10.1017/S0142716419000328

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. doi:10.1037/0033-2909.131.1.66

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. doi:10.1207/s15327868ms0304_1

Kazmerski, V. A., Blasko, D. G., & Dessalegn, B. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory & Cognition*, *31*, 673–689. doi:10.3758/BF03196107

Keane, M., & Costello, F. (2001). Setting limits on analogy: Why conceptual combination is not structural alignment. In D. Gentner, K. J. Holyoak, & B. Kokinov (Eds.), *The analogical minds: A cognitive science perspective* (pp. 287–312). Cambridge, MA: MIT Press.

Kenett, Y. N., Gold, R., & Faust, M. (2018). Metaphor comprehension in low and high creative individuals. *Frontiers in Psychology*, *9*, 482. doi:10.3389/fpsyg.2018.00482

Keysar, B., Shen, Y., Glucksberg, S., & Horton, W. S. (2000). Conventional language: How metaphorical is it? *Journal of Memory and Language*, *43*, 576–593. doi:10.1006/jmla.2000.2711

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*, 257–266. doi:10.3758/BF03212981

Kintsch, W. (2001). Predication. *Cognitive Science*, *25*, 173–202. doi:10.1207/s15516709cog2502_1

Kintsch, W., & Bowles, A. R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, *17*, 249–262. doi:10.1207/S15327868MS1704_1

Kintsch, W., & Mangalath, P. (2011). The construction of meaning. *Topics in Cognitive Science*, *3*, 346–370. doi:10.1111/j.1756-8765.2010.01107.x

Kintsch, W., & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. doi:10.1037/0033-295X.85.5.363

Kövecses, Z. (2015). *Where metaphors come from: Reconsidering context in metaphor*. Oxford: Oxford University Press.

Kubricht, J. R., Lu, H., & Holyoak, K. J. (2017). Individual differences in spontaneous analogical transfer. *Memory & Cognition*, *45*, 576–588. doi:10.3758/s13421-016-0687-7

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago, IL: University of Chicago Press.

Lemaire, B., & Bianco, M. (2003). Contextual effects on metaphor comprehension: Experiment and simulation. In F. Detje, D. Dörner, & H. Schaub (Eds.), *Proceedings of the 5th international conference on cognitive modelling* (pp. 153–158). Bamberg, Germany: Universitäts-Verlag.

Li, H., Zhu, K. Q., & Wang, H. (2013). Data-driven metaphor recognition and explanation. *Transactions of the Association for Computational Linguistics*, *1*, 379–390. doi:10.1162/tacl_a_00235

McGlone, M. S., & Manfredi, D. A. (2001). Topic-vehicle interaction in metaphor comprehension. *Memory & Cognition*, *29*, 1209–1219. doi:10.3758/BF03206390

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, *26*, 3111–3119.

Morsanyi, K., Stamenković, D., & Holyoak, K. J. (2020b). Analogical reasoning in autism: A systematic review and meta-analysis. In K. Morsanyi & R. Byrne (Eds.), *Thinking, reasoning, and decision making in autism* (pp. 59–87). London, UK: Routledge. doi:10.4324/9781351060912-4

Morsanyi, K., Stamenković, D., & Holyoak, K. J. (2020a). Metaphor processing in autism: A systematic review and meta-analysis. *Developmental Review*, *57*, 100925. doi:10.1016/j.dr.2020.100925

Mulholland, T. M., Pellegrino, J. W., & Glaser, R. (1980). Components of geometric analogy solution. *Cognitive Psychology*, *12*, 252–284. doi:10.1016/0010-0285(80)90011-0

Nayak, N. P., & Gibbs, R. W., Jr. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology: General*, *119*, 315–330. doi:10.1037/0096-3445.119.3.315

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 398–415. doi:10.1037/0278-7393.17.3.398

Olkoniemi, H., Ranta, H., & Kaakinen, J. H. (2016). Individual differences in the processing of written sarcasm and metaphor: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 433–450. doi:10.1037/xlm0000176

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. L. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, *17*, 465–477. doi:10.1016/S0022-5371(78)90283-9

Pexman, P. M., Ferretti, T. R., & Katz, A. N. (2000). Discourse factors that influence on-line reading of metaphor and irony. *Discourse Processes*, *29*, 201–222. doi:10.1207/S15326950dp2903_2

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. doi:10.1080/10926480802568422

Prat, C. S., Mason, R. A., & Just, M. A. (2012). An fMRI investigation of analogical mapping in metaphor comprehension: The influence of context and individual cognitive capacities on processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 282–294. doi:10.1037/a0026037

Rai, S., & Chakraverty, S. (2020). A survey on computational metaphor processing. *ACM Computing Surveys*, *53*(2), Article 24. doi:10.1145/3373265

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. doi:10.1016/j.neuroimage.2012.06.022

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

Roncero, C., & de Almeida, R. G. (2015). Semantic properties, aptness, familiarity, conventionality, and interpretive diversity scores for 84 metaphors and similes. *Behavior Research Methods*, *47*, 800–812. doi:10.3758/s13428-014-0502-y

Rumelhart, D. (1979/1993). Some problems with the notion of literal meanings. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 71–82). Cambridge, UK: Cambridge University Press.

Salthouse, T. A. (2004). Localizing age-related individual differences in a hierarchical structure. *Intelligence*, *32*(6). doi:10.1016/j.intell.2004.07.003

Shinjo, M., & Myers, J. L. (1987). The role of context in metaphor comprehension. *Journal of Memory and Language*, *26*, 226–241. doi:10.1016/0749-596X(87)90125-2

Silvia, P. J., & Beaty, R. E. (2012). Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence*, *40*(4), 343–351. doi:10.1016/j.intell.2012.02.005

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.

Stamenković, D., Ichien, N., & Holyoak, K. (2019). Metaphor comprehension: An individual-differences approach. *Journal of Memory and Language*, *105*, 108–118. doi:10.1016/j.jml.2018.12.003

Steen, G. (1994). *Understanding metaphor in literature: An empirical approach*. New York, NY: Longman Publishing.

Stern, J. (2000). *Metaphor in context*. Cambridge, MA: MIT Press.

Thibodeau, P. H., & Durgin, F. H. (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. *Journal of Memory and Language*, *58*, 521–540. doi:10.1016/j.jml.2007.05.001

Tourangeau, R., & Sternberg, R. J. (1981). Aptness in metaphor. *Cognitive Psychology*, *13*, 27–55. doi:10.1016/0010-0285(81)90003-7

Tourangeau, R., & Sternberg, R. J. (1982). Understanding and appreciating metaphors. *Cognition*, *11*, 203–244. doi:10.1016/0010-0277(82)90016-6

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. doi:10.1207/s15327868ms0103_3

van Gogh, V. (1882). *Letter to Theo van Gogh*. Retrieved from http://www.webexhibits.org/vangogh/letter/11/253.htm

Vartanian, O. (2012). Dissociable neural systems for analogy and metaphor: Implications for the neuroscience of creativity. *British Journal of Psychology*, *103*, 302–316. doi:10.1111/j.2044-8295.2011.02073.x

Vendetti, M., Wu, A., & Holyoak, K. J. (2014). Far out thinking: Generating solutions to distant analogies promotes relational thinking. *Psychological Science*, *25*, 928–933. doi:10.1177/0956797613518079

Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, inhibition and analogical reasoning in older adults. *Psychology and Aging*, *19*, 581–591. doi:10.1037/0882-7974.19.4.581

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition*, *28*, 1205–1212. doi:10.3758/BF03211821

Wechsler, D. (1997). *WAIS-III: Administration and scoring manual – Wechsler adult intelligence scale*. London, UK: The Psychological Corporation/Pearson Education.

Wheelwright, P. (1968). *The burning fountain: A study in the language of symbolism*. Bloomington: Indiana University Press.

Wilson, D., & Carston, R. (2006). Metaphor, relevance and the 'emergent property' issue. *Mind & Language*, *21*, 404–433. doi:10.1111/j.1468-0017.2006.00284.x

Wisniewski, E. J., & Gentner, D. (1991). On the combinatorial semantics of noun pairs: Minor and major adjustments to meaning. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 241–284). Amsterdam, the Netherlands: North Holland.