An Individual-Differences Approach to Poetic Metaphor: Impact of Aptness and Familiarity

Dušan Stamenković, Katarina Milenković, Nicholas Ichien & Keith J. Holyoak

Published online: 16 Mar 2023.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

An Individual-Differences Approach to Poetic Metaphor: Impact of Aptness and Familiarity

Dušan Stamenković [a,d], Katarina Milenković [b], Nicholas Ichien [c], and Keith J. Holyoak [c]

aSchool of Culture and Education, Södertörn University, Huddinge, Sweden; bFaculty of Sciences and Mathematics, University of Niš, Serbia; cDepartment of Psychology, University of California, Los Angeles, USA; dFaculty of Philosophy, University of Niš, Serbia

ABSTRACT

Using poetic metaphors in the Serbian language, we identified systematic variations in the impact of fluid and crystalized intelligence on comprehension of metaphors that varied in rated aptness and familiarity. Overall, comprehension scores were higher for metaphors that were high rather than low in aptness, and high rather than low in familiarity. A measure of crystalized intelligence was a robust predictor of comprehension across the full range of metaphors, but especially for those that were either relatively unfamiliar or more apt. In contrast, individual differences associated with fluid intelligence were clearly found only for metaphors that were low in aptness. Superior verbal knowledge appears to be particularly important when trying to find meaning in novel metaphorical expressions, and also when exploring the rich interpretive potential of apt metaphors. The broad role of crystalized intelligence in metaphor comprehension is consistent with the view that metaphors are largely understood using semantic integration processes continuous with those that operate in understanding literal language.

Introduction

A metaphor aims to provide insight into one thing by describing it as another. Particularly for literary metaphors (often drawn from poetry), the insight being expressed may be multi-facetted and open-ended (Wheelwright, 1968). In "Consciousness of a Poem," the Serbian poet Branko Miljković (1956/2005) metaphorically describes "blood" as "powerless ink." Here the *target* concept (what is being talked about) is "blood," and the *source* (used to characterize the target) is "powerless ink." The metaphor is supported by the fact that blood and ink are both types of liquid; but of course, blood is not literally ink, and the metaphor does not mean "blood is a poor type of ink." Readers may arrive at a variety of interpretations. For example, ink has a symbolic link (via metonymy) to pen and to writing, and might call to mind the claim that, "The pen is mightier than the sword." Blood is linked (again via metonymy) to physical violence (the effect of the sword), and the shedding of one's own blood can be a form of extreme sacrifice. To say that "blood is powerless ink" might therefore mean that sacrificing one's life is less effective (perhaps in promoting a cause) than is writing. Perhaps the poet is saying something like, "sacrificing one's life is not the most effective means of achieving our purpose."

Metaphor plays a particularly important role in poetry (Holyoak, 2019; Lakoff & Turner, 1989); however, within the body of research on metaphor, empirical studies involving poetic metaphors have been relatively rare (e.g., Gibbs & Bogdonovich, 1999; Marks, 1982; Stamenković, Ichien, & Holyoak, 2019a, 2020; Tourangeau & Rips, 1991; for a recent review see Glicksohn & Goodblatt, 2021). Here we

---

CONTACT Dušan Stamenković ✉ dusan.stamenkovic@filfak.ni.ac.rs ◻ Language Cognition Laboratory, Faculty of Philosophy, University of Niš, Kabinet 418, Ćirila i Metodija 2, Niš 18000, Serbia

report a study of the comprehension of poetic metaphors that vary in both *aptness* (quality of the metaphoric description) and *familiarity*. We examine individual differences in ability to comprehend poetic metaphors that relate to *fluid intelligence* (the ability to reason about novel problems) and also *crystalized intelligence* (verbal knowledge accumulated over a person's lifetime).

Although the 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 measure of the unexpectedness of words), relative dissimilarity of source and target concepts, and the combination of concrete words with relatively complex grammar and high lexical diversity (see also Baggio, 2018). These differences suggest that literary metaphors tend to be high in general cognitive complexity.

Both the individual differences (in fluid and crystalized intelligence) and metaphor properties (aptness and familiarity) investigated in the present study of poetic metaphors have figured prominently in general debates about the cognitive mechanisms involved in metaphor comprehension (for a review see Holyoak & Stamenković, 2018). One general hypothesis is that metaphor comprehension requires *analogical reasoning* to relate the target to the source (Gentner & Clement, 1988; Tourangeau & Rips, 1991; Tourangeau & Sternberg, 1981). Analogical reasoning is generally viewed as a process of comparing structured representations of the source and target, and then using the source to comprehend the less-familiar target (Holyoak, 2012). An alternative view assumes that metaphor processing is continuous with literal language comprehension, involving *semantic integration*: the incremental construction of the meaning of a message from the meanings of words, guided by both syntactic structure and the pragmatic context of the utterance (Kintsch & Mangalath, 2011; see Glucksberg & Keysar, 1990, for the more specific hypothesis that metaphors are based on categorization). Whereas analogical reasoning is typically viewed as operating on complex knowledge structures held in working memory, semantic integration may be based in part on computationally simpler processes (e.g., operations on vectors of semantic features) that operate at the level of lexical semantics (Kintsch, 2000).

In addition to focusing on literary metaphors, we wished to expand the empirical investigation of metaphor to languages other than English. Studies of non-English metaphors have been infrequent (but see Aisenman, 1999; Cacciari & Glucksberg, 1995; Utsumi, 2007), and to our knowledge none have specifically examined poetic metaphors. The present study involved metaphors drawn from Serbian poetry, administered to Serbian speakers.

Individual differences in metaphor comprehension

Classical theories of intelligence (Cattell, 1971) distinguish between 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. Crystalized intelligence involves reasoning (typically verbal) that draws upon prior knowledge, and hence is likely to support semantic integration. There is evidence that both of these forms of intelligence impact metaphor comprehension. For example, Chiappe and Chiappe (2007) found that individuals who scored high on a working-memory test (linked to fluid intelligence) generated higher-quality interpretations of metaphors more quickly. In addition, these investigators found that measures of vocabulary knowledge and exposure to printed text (linked to crystalized intelligence) predicted quality of created metaphors. Indeed, crystalized intelligence yielded somewhat higher correlations with metaphor interpretation and production than did measures of working memory. Using a task involving cued production of metaphors, Beaty and Silvia (2013) found that fluid intelligence was strongly related to production of creative novel metaphors, whereas crystalized knowledge predicted the ability to generate apt conventional metaphors. Overall, these findings suggest both fluid and crystalized intelligence have an impact on metaphor processing, with varying importance of one or the other and depending on the type of metaphor or nature of the task.

Stamenković et al. (2019a) applied an individual-differences approach to examine the comprehension of metaphors drawn from literary as well as nonliterary sources. They argued that because good literary metaphors are high in cognitive complexity, they may be more likely to elicit analogical reasoning, which is known to place high demands on fluid intelligence (e.g., Cho, Holyoak, & Cannon, 2007). Using regression analyses, Stamenković et al. (2019a) found that for nonliterary metaphors, only measures of crystalized intelligence predicted unique variance in comprehension scores. For literary metaphors, in contrast, measures of fluid and crystalized intelligence each made separable contributions to predicting comprehension. In a subsequent study, Stamenković et al. (2020) examined individual differences in comprehending literary metaphors as a function of variations in a verbal context. Measures of fluid and crystalized intelligence both made separable contributions to predicting metaphor comprehension, with a supportive context increasing the contribution of crystalized verbal intelligence.

The general picture from these studies is that crystalized intelligence impacts comprehension across all metaphor types, whereas fluid intelligence (on which analogical reasoning depends heavily) plays a greater role for more complex literary metaphors, especially when presented without a supportive context. The lesser contribution of fluid intelligence for nonliterary metaphors is consistent with a dissociation that has been noted between metaphor comprehension and analogical reasoning for people with autism spectrum disorder, who exhibit impaired metaphor processing despite spared ability in analogical reasoning (Morsanyi, Stamenković, & Holyoak, 2020a, 2020b). Evidence from neuroimaging also suggests that metaphor processing does not primarily depend on brain regions associated with analogical reasoning (for meta-analyses see Bohrn, Altmann, & Jacobs, 2012; Rapp, Mutschler, & Erb, 2012; Vartanian, 2012).

Aptness and familiarity

Many studies of metaphor processing have examined the impact of variations in aptness and familiarity (as well as the related factor of conventionality). Aptness is high when the source is perceived as providing a unique and accurate description of the target, such that salient properties of the source are attributed to the target (Al-Azary & Katz, 2021; Blasko & Connine, 1993; Chiappe, Kennedy, & Smykowski, 2003b; Gagné, 2002; Jones & Estes, 2005, 2006; Roncero & de Almeida, 2015). More apt metaphors are processed faster and are more easily understood (Blasko & Connine, 1993; Chiappe & Kennedy, 1999; Chiappe, Kennedy, & Chiappe, 2003a; Chiappe et al., 2003b; Glucksberg & McGlone, 1999; Jones & Estes, 2005, 2006).

Whereas aptness is based on how well the source is perceived to "fit" the target, familiarity and conventionality are based on frequency of exposure to either the full metaphor (familiarity) or to the source used as a metaphor (conventionality). Familiarity with a given metaphor will presumably reduce processing difficulty, consistent with the general observation that practice leads to faster processing; and studies have found that more familiar metaphors are indeed read faster than less familiar ones (e.g., Blasko & Briihl, 1997; Blasko & Connine, 1993). Bowdle and Gentner (2005) proposed that a novel metaphor is comprehended using analogy, but after repeated use comes to be understood as a category statement. However, a number of studies have found that aptness is a good predictor of various measures of metaphor processing, in some cases even better than conventionality or familiarity (Chiappe et al., 2003a, 2003b; Jones & Estes, 2006; Wong & Holyoak, 2021). In particular, aptness rather than conventionality seems to predict the extent to which a metaphor is interpreted as an expression of categorization (Jones & Estes, 2005).

In the present study we examined how measures of fluid and crystalized intelligence predict comprehension of poetic metaphors that have been normed for both aptness and familiarity.

Method

Participants

A total of 107 undergraduate and graduate students at the University of Niš, Serbia participated in the study for course credit. This sample size is comparable to those in similar studies of individual differences in metaphor comprehension (Stamenković et al., 2019a, 2020). The study was approved by the Research Ethics Committee of the Faculty of Philosophy, University of Niš. The participants were all native speakers of Serbian. One participant failed to complete the entire battery of tests, and data from 11 participants were dropped from analyses based on the same set of criteria used in our previous studies (criteria indicative of carelessness or inattention on the verbal tasks): score of below 12 (max = 40) on the Semantic Similarities Test and/or below 21 (max = 84) on the metaphor test. After these exclusions, we report results for a total of 95 participants (female = 68, male = 27, mean age = 21.6 years): 49 participants who were randomly assigned to a test set in which metaphors were selected to strongly vary on aptness, and 46 to a test set in which metaphors were selected to strongly vary on familiarity.

Design, materials, and procedure

All participants completed three tasks in a fixed order. The first two were assessments of individual differences on cognitive tasks, and the third and final task involved metaphor comprehension. All tasks were administered to participants individually using Google Forms. None of the tasks involved any time pressure. The entire test session lasted approximately 50 minutes.

Task 1: raven's progressive matrices (RPM)

A short version (Arthur, Tubre, Paul, & Sanchez-Ku, 1999) was administered to assess fluid intelligence.

Task 2: semantic similarities test (SST)

Crystalized verbal intelligence was assessed using the Serbian version of the SST (available in the Supplemental Online Material). The SST (Stamenković et al., 2019a) is designed to measure ability to identify similarities between concepts expressed as pairs of single words, where the similarities vary in degree of abstraction. The test is similar to the Similarities subscale of the Wechsler Adult Intelligence Scale (WAIS). In addition to predicting comprehension of both literary and nonliterary metaphors (Stamenković et al., 2019a, 2020), the SST has been found to predict performance on a battery of relational reasoning tasks (Gray & Holyoak, 2020). In the test, word pairs are presented with the question, "How are the two concepts in each pair similar to one another?" Scores on the SST correlate well with the Vocabulary subscale of the Wechsler Adult Intelligence Scale, a standard measure of crystalized intelligence (Stamenković et al., 2019a, 2020).

Task 3: poetic metaphor comprehension

The final task consisted of a set of 28 literary metaphor comprehension items, selected from a list of literary metaphors drawn from Serbian poetry and normed on several dimensions/features (Stamenković, Milenković, & Dinčić, Stamenković, et al., 2019b). The norming study (largely based on the procedures used in Katz, Paivio, Marschark, & Clark, 1988) included 55 poetic metaphors selected by a literary expert from over 65 nineteenth-century and twentieth-century poems written by a range of Serbian poets, including Branko Radičević, Laza Kostić, Vojislav Ilić, Đura Jakšić, Desanka Maksimović, and Branko Miljković. The poems selected for the norming study were intended to reflect the diversity of poetic movements and styles across these two centuries. The expert had the task of extracting all metaphorical expressions from these poems. These were then grouped, with all similar/duplicate metaphors counted as one. All metaphors were then converted to the *<nominal> is <nominal>* form, yielding the final list of 55 items.

In the norming study conducted by Stamenković et al. (2019b), the 55 metaphors were rated for quality, metaphoricity, aptness, familiarity, comprehensibility and source-target similarity by 235 Serbian-speaking participants using a 7-point Likert scale (min = 1, max = 7). Participants in the norming study had diverse profiles in terms of primary, secondary and tertiary education, and doubtless varied in level of overall poetry appreciation; however, participants were unlikely to have encountered many of the poems included in the pool. A large majority of the selected poems are not included in the elementary or high-school curricula in Serbia, so these levels of formal education did not serve as a source of familiarity. The interscale correlations for Serbian poetic metaphors reported by Stamenković et al. (2019b) tended to be lower than those obtained with literary metaphors in English (Katz et al., 1988), but the pattern of correlations was generally similar.

For the present study, items were selected based on ratings of *aptness* and *familiarity*. The question used by Stamenković et al. (2019b) to assess aptness was, "In the expression <x>, how apt is term A < source> for describing term B < target>"; the question used to assess familiarity was, "How familiar is the following metaphorical expression to you?" While these two ratings were significantly correlated (*r* (40) = .59, *p* < .001), we will report an analysis that separates the contributions of a metaphor's aptness versus familiarity as predictors of the ease with which human reasoners interpret that metaphor (see section *Individual Differences in Metaphor Comprehension in Relation to Aptness and Familiarity*). Specifically, we will examine the relation between metaphor comprehension and each variable (e.g., aptness) with the other variable (e.g., familiarity) equated.

For the present study, two test sets were created by selecting subsets of the normed metaphors based on maximal differences in either aptness or familiarity. Each test was designed to create a split-half variable, consisting of a total of 28 items of which 14 were those rated lowest and 14 those rated highest on the relevant rating scale. Because the two ratings were moderately correlated, this selection procedure created sets that partially overlapped (e.g., *The soul is a heavenly star* was rated high on both aptness and familiarity, whereas *A man is a crumbling boat* was rated low on both scales). Of the total of 42 unique items used across the two tests, 14 items were unique to one set and 14 occurred in both sets. Each metaphor was in the form of a single critical sentence (e.g., *The soul is a narrow dungeon*). Their syntactic forms included nominal (*<nominal> is <nominal>*, e.g., *Love is radiance*), nominal with an adjective modifier (e.g., *Life is a poisonous flower*), and nominal with a prepositional phrase (e.g., *The sun is the witness of time*). Examples of metaphors belonging to each set are given in Table 1. All the poetic metaphors used in the present study, and their English translations, are provided in Appendix I (Supplemental Online Material).

For the metaphor task, participants were randomly assigned to one of two groups, each of which received one of the two items sets (based on variation in aptness for else familiarity). The metaphor task was otherwise identical for the two conditions. Metaphors were presented one at a time in a randomized order for each participant. Comprehension was assessed by an open-ended question asking participants to type their interpretation of each metaphorical statement.

Results

Metaphor task rating and coding

Responses to the poetic metaphor comprehension task were scored by two independent raters (both linguists and both native speakers of Serbian) using a 4-point scale, with responses being awarded 0, 1, 2, or 3 points. A score of 3 (completely plausible) 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 and showed full understanding of the metaphor). A score of 2 (mostly plausible) was given if the paraphrase described the metaphorical meaning, but maintained explicit links to the source domain and focused mostly on one domain. A score of 1 (partly plausible) was given if the paraphrase was strongly linked to the source domain and remained relatively simple, or if it ascribed physical features to the target. A score of 0 (implausible) was given if the paraphrase

Table 1. Examples of items from the two test sets used to assess metaphor comprehension(with mean Aptness and familiarity ratings).

Aptness Set	
High Aptness	**Low Aptness**
(mean aptness 4.85, SD = .45)	(mean aptness 2.51, SD = .19)
(mean familiarity 3.52, SD = .68)	(mean familiarity 2.41, SD = .45)
The soul is a precious stone.	*A man is a tomb.*
The moon is a dark night's guard.	*Life is a pitiful shadow.*
An eye is a sea.	*A man is a barren ash field.*
Familiarity Set	
High Familiarity	**Low Familiarity**
(mean familiarity 3.84, SD = .30)	(mean familiarity 2.02, SD = .21)
(mean aptness 4.40, SD = .92)	(mean aptness 3.02, SD = .44)
The sky is an ocean.	*An umbrella is a black sail.*
A man is a rock.	*A waterfall is a wild, unbridled horse.*
A woman is a flower.	*A tomb is the flower of glory.*

restated the metaphor using simpler terms with no further insights (i.e., was literal in nature), if it was a complete misinterpretation or nonsensical, or if no response was entered. Examples of item scoring are provided in Appendix II (Supplemental Online Material). Cohen's κ was computed to assess degree of agreement between the two raters' judgments of the quality of the metaphor interpretations. The agreement between the judgments of the two raters was adequate, κ = .78, p < .001. In case of a disagreement, after a discussion, in most cases the higher of the two scores was assigned.

Individual differences in comprehension of metaphors

Table 2 summarizes the descriptive statistics and intercorrelations among the three measures we obtained (RPM, SST, and poetic metaphor comprehension) for both item sets. As in previous studies (Gray & Holyoak, 2020; Stamenković et al., 2019a, 2020), there was a reliable correlation between scores on the RPM and SST tests, $r(140)$ = .36, p < .001, based on all participants.

We examined the impact of individual differences in fluid and crystalized verbal intelligence on metaphor comprehension scores, separately for low and high item subsets within the aptness and familiarity sets. Figure 1 shows our individual difference measures plotted against metaphor comprehension scores for each subset. Correlation and regression analyses were performed to assess the interrelationships among the RPM, SST and mean score on metaphor comprehension tests. Table 3 presents both the raw correlations of each of the two predictor variables with metaphor comprehension scores, and also partial correlations obtained from the regression analysis. The regression analyses revealed that for all four item sets, SST scores predicted separable variance in comprehension

Table 2. Descriptive statistics for each measure for each group.

| Group 1: Aptness Ratings Varied				
Test	N	Mean	SD	Range
RPM	49	6.63	2.86	1–11
SST	49	29.63	3.87	20–38
Comprehension: High Aptness (Mean)	49	1.88	.38	1.07–2.71
Comprehension: Low Aptness (Mean)	49	1.61	.42	.64–2.79
Group 2: Familiarity Ratings Varied				
Test	N	Mean	SD	Range
RPM	46	6.24	2.62	1–11
SST	46	28.41	3.72	18–34
Comprehension: High Familiarity (Mean)	46	1.83	.50	.46–2.64
Comprehension: Low Familiarity (Mean)	46	1.24	.46	.36–2.29
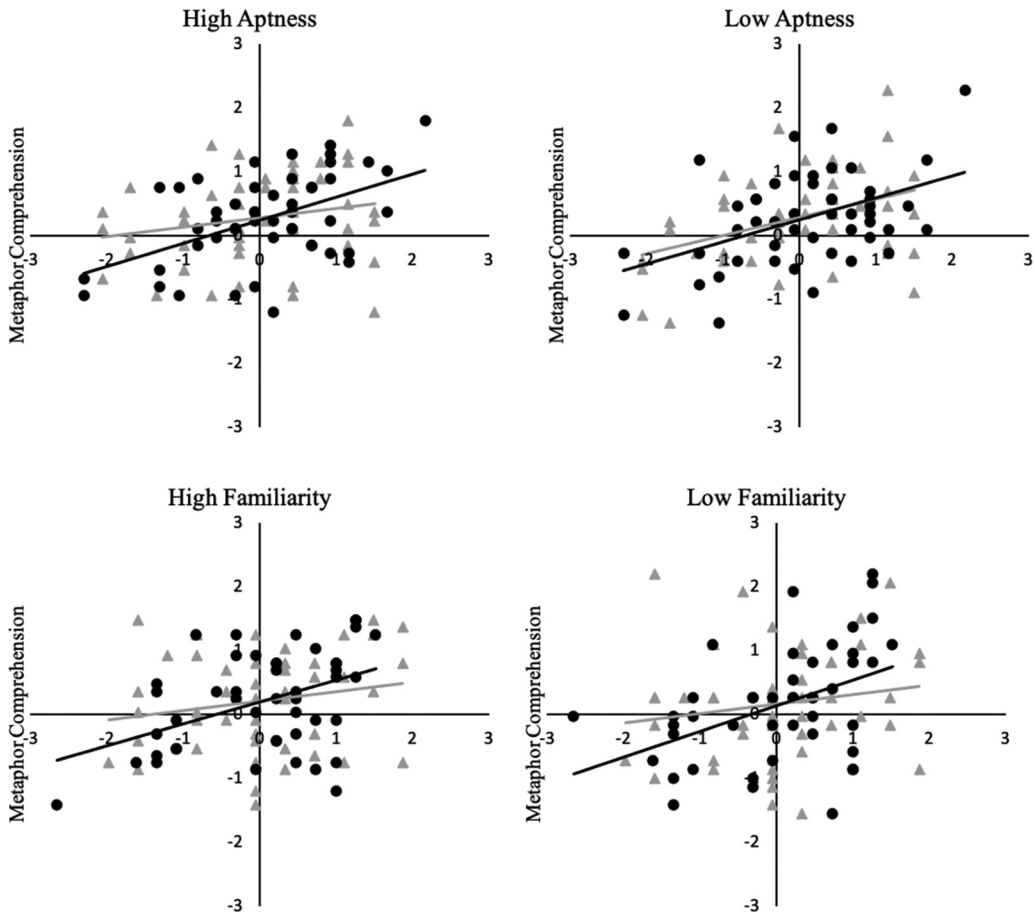
Figure 1. Individual differences in fluid intelligence (standardized RPM scores) and crystalized verbal intelligence (standardized SST scores) plotted against standardized metaphor comprehension scores (▲ RPM ● SST).

Table 3. Correlations and partial correlations of individual-difference measures with metaphor comprehension scores for each item subset.

		RPM		SST	
Group		Raw	Partial	Raw	Partial
Aptness	High Aptness	.21	.05	.49***	.457**
	Low Aptness	.40**	.29*	.46***	.37**
Familiarity	High Familiarity	.19	.04	.43**	.39**
	Low Familiarity	.17	.01	.44**	.40**

Note: * $p < .05$, ** $p < .01$, *** $p < .001$

accuracy, with partial correlations ranging from .37 (SST for metaphors low in aptness, $p < .01$) to .46 (SST for metaphors high in aptness, $p < .01$). In contrast, RPM scores correlated only with the comprehension of metaphors low in aptness, with a raw correlation of .40 ($p < .01$) and a weaker partial correlation of .29 ($p < .05$). This pattern suggests that while both measures have an impact on metaphor comprehension, for apt poetic metaphors SST (crystalized intelligence) is more important than is RPM (fluid intelligence).

Individual differences in metaphor comprehension in relation to aptness and familiarity

We next examined the joint influence on metaphor comprehension scores of metaphor aptness and familiarity coupled with individual differences in fluid and crystalized verbal intelligence. For this analysis, we combined the responses to the two versions of the metaphor test, including both aptness and familiarity ratings as predictors for each individual item. We used the *clmm* function from version 2019.12.10 of the ordinal package (Christensen, 2019) in R version 3.6.2 (R Core Team, 2017) to define and fit a cumulative link mixed-effect model of metaphor comprehension scores. This approach allowed us to tease apart the relative contributions not only of fluid and crystallized intelligence, as reported above, but also of metaphor aptness and familiarity on metaphor comprehension. Our full model included participant and metaphor item as random effects, with aptness, familiarity, RPM scores, and SST scores as mean-centered fixed effects. In order to test whether either aptness and familiarity altered the relationship between our two individual difference measures and metaphor comprehension scores, our full model also included two three-way interaction terms, aptness*RPM*SST and familiarity*RPM*SST, as predictors specified as fixed effects. An analysis of deviance (ANODE) between the full model and an otherwise equivalent model without interaction terms revealed that the addition of interaction terms resulted in the full model accounting for more variance in comprehension scores than did the reduced model (ΔAIC = −1.3, LLR $\chi^2(7)$ = 15.31, p = .03). Table S1 (in the Supplemental Online Material) provides a summary of the full model.

A significant three-way interaction relating aptness to the two individual difference measures (β = .11, z = 2.25, p = .025) indicated that aptness changed the relationship between metaphor comprehension scores and RPM and SST. Moreover, significant two-way interactions between aptness and SST (β = .15, z = 2.75, p = .006) and between aptness and RPM (β = −.10, z = −2.06, p = .039), indicate that holding RPM score at its mean value, an increase in SST benefits comprehension on *more* apt metaphors, whereas holding SST score at its mean value, an increase in RPM score benefits comprehension on *less* apt metaphors. Figure 2 presents interaction plots of estimated marginal means based on the full model. These plots illustrate that whereas higher scores on the RPM consistently predict better comprehension for *less* apt metaphors, it is only with concurrently high scores on the SST that high RPM scores predicts better comprehension for *more* apt metaphors. It thus seems that stronger executive function (assessed by RPM) aids a reasoner in finding meaning in the somewhat questionable comparisons characteristic of less apt metaphors; whereas superior verbal knowledge (assessed by the SST) is necessary in order for stronger executive function to aid in finding meaning in the deeper comparisons characteristic of more apt metaphors.

The three-way interaction effect between familiarity, RPM, and SST was not significant (β = .04, z = 0.67, p = .500). However, a significant two-way interaction between familiarity and SST (β = −.13, z = −2.48, p = .013) indicates that holding RPM score at its mean value, an increase in SST score benefits comprehension of *less* familiar metaphors. Reasoners with superior verbal knowledge thus appear to have greater facility in finding meaning in novel metaphorical expressions.

Our full model also revealed main effects of SST (β = .40, z = 4.72, p < .001) and familiarity (β = .44, z = 3.01, p = .002), indicating that each of these variables predicted variations in metaphor comprehension scores when all other predictors were set to the sample mean. Consistent with the findings of Stamenković et al. (2019a, 2020), individual differences in crystalized intelligence (assessed by SST) predicted metaphor comprehension across a wider range of metaphor types than did fluid intelligence (assessed by RPM).
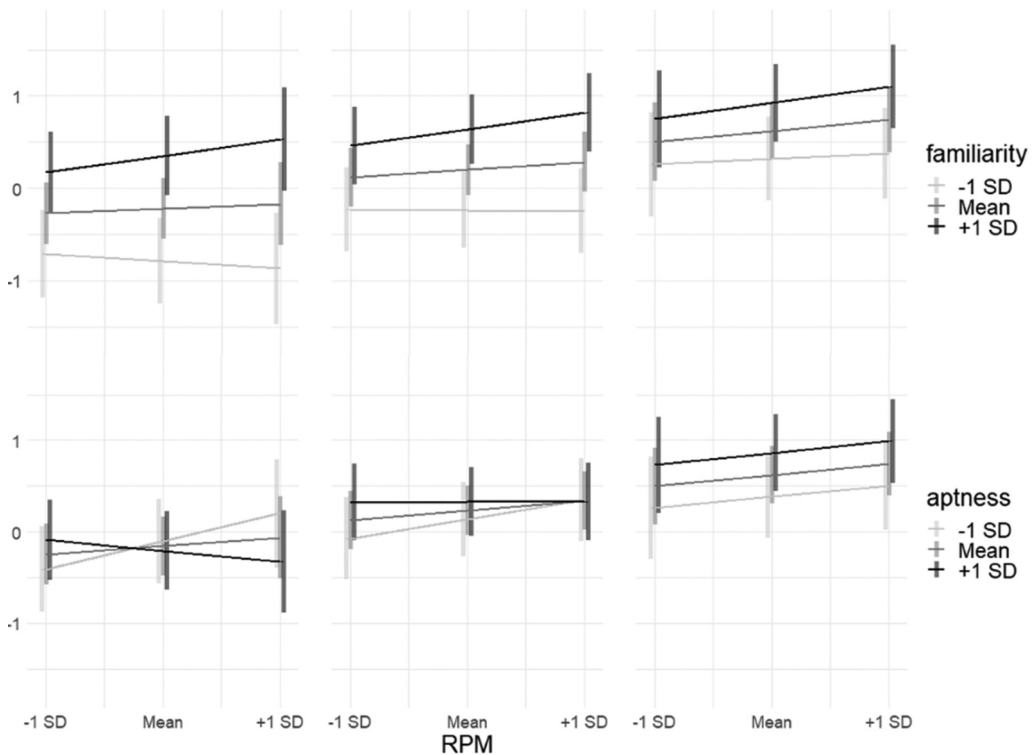
Figure 2. Interaction plots depicting estimated marginal means of log odds based on an ordinal mixed-effects regression model. Plots in the left column show estimates with SST score set to 1 standard deviation below the sample mean; those in the middle column, at the sample mean; and those in the right column, 1 standard deviation above the sample mean. Error bars reflect 95% confidence intervals of model estimates.

Discussion

In a study of the comprehension of literary metaphors from Serbian poetry (with Serbian participants), we identified systematic variations in the impact of fluid and crystalized intelligence across metaphors that varied in rated aptness and familiarity. Overall, comprehension scores were higher for metaphors that were high rather than low in aptness, and high rather than low in familiarity. However, these variables interacted with fluid intelligence (measured by the Ravens Progressive Matrices test) and crystalized intelligence (measured by the Semantic Similarities Test). Consistent with previous findings for English metaphors (Stamenković et al., 2019a, 2020), the measure of crystalized intelligence was a robust predictor of comprehension across the full range of metaphors, but especially for those that were either relatively unfamiliar or more apt. Superior verbal knowledge appears to be particularly important when trying to find meaning in novel metaphorical expressions, and also when exploring the rich interpretive potential of apt metaphors. Because crystalized intelligence is directly linked to knowledge of lexical semantics, its broad role in metaphor comprehension is consistent with the view that metaphors are understood using semantic integration processes continuous with those that operate in understanding literal language (Kintsch, 2000).

In contrast, individual differences associated with fluid intelligence were clearly found only for metaphors that were low in aptness. For more apt metaphors, fluid intelligence conveyed an advantage only for those participants who also scored high in crystalized intelligence. This pattern suggests that explicit analogical reasoning (which draws heavily on fluid intelligence) may provide a back-up

strategy to help find meaning in the relatively weak comparisons characteristic of less apt metaphors. In addition, for people with superior verbal knowledge, analogical reasoning may augment the ability to find meaning in the deeper comparisons provided by apt metaphors.

Future work should examine in more detail the nature of aptness as a quality of metaphors. Particularly in poetic metaphors, the "point" is not simply to provide an informative description of the target, but to generate an esthetic experience that includes emotional impact. The emotional impact of poetry is linked to concrete language that conveys sensory and motoric detail (Holyoak, 2019). A study of neural responses during the act of writing a poem found that when experts generated poems rated high in linguistic creativity, subregions of the frontal cortex associated with cognitive control were less closely coupled with areas that support sensory and motor activity (Liu et al., 2015). For poets, the control network apparently reduces its regulation of these areas, allowing sensory imagery to become available so that it finds its way into the emerging poem. When a poetic metaphor is comprehended, it may seem especially apt if it succeeds in eliciting sensorimotor associations (sometimes accompanied by visual imagery, or by a kind of internal simulation of bodily responses; see Jamrozik, McQuire, Cardillo, & Chatterjee, 2016). There is evidence that sensorimotor associations are more likely to be evoked by metaphors that are less familiar or conventional (Al-Azary & Katz, 2021). Novelty may play a role in determining esthetic responses to metaphors, thereby contributing to their perceived aptness. Future research should also examine in more detail the variables of surprisal, semantic distance between source and target concepts, comprehension difficulty, grammatical and lexical complexity and diversity, and their relations to and interactions with aptness.

The pattern of results observed in the present study using the Serbian language – particularly the dominant role of crystalized intelligence in comprehending metaphors, including those drawn from literary sources – is consistent with findings from previous studies of metaphor comprehension involving native speakers of English (Stamenković et al., 2019a, 2020). This convergence is encouraging, in that it supports the view that the psychological mechanisms underlying metaphor processing may be fundamentally the same across human languages. It is also noteworthy that similar patterns of inter-scale and inter-dimension correlations were found by Katz et al. (1988) for English metaphors and Stamenković et al. (Stamenković, et al., 2019b) for Serbian metaphors. Future studies will be needed to determine whether this convergence extends across a broader range of languages. In addition, future studies should examine the impact of additional individual differences that may relate to metaphor processing, such as those involving creative potential (e.g., Kenett, Gold, & Faust, 2018), measures of inhibitory control (Pierce & Chiappe, 2008; Sana, Park, Gagné, & Spalding, 2021), as well as variations in personality traits, preference for imagery, and emotional understanding (see Fetterman, Bair, Werth, Landkammer, & Robinson, 2016).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Open practices statement

The data, the materials for all experiments and Supplemental Online Material are available in Mendeley Data (http://dx.doi.org/10.17632/8wfj3285gd.1). The study was approved by the Research Ethics Committee of the Faculty of Philosophy, University of Niš. The study was not preregistered.

Funding

ORCID

Dušan Stamenković http://orcid.org/0000-0002-0121-4591
Katarina Milenković http://orcid.org/0000-0001-8188-858X
Nicholas Ichien http://orcid.org/0000-0002-0928-0809
Keith J. Holyoak http://orcid.org/0000-0001-8010-6267

References

Aisenman, R. A. (1999). Structure-mapping and the simile-metaphor preference. *Metaphor and Symbol*, 14(1), 45–51. doi:10.1207/s15327868ms1401_5

Al-Azary, H., & Katz, A. N. (2021). Do metaphorical sharks bite? Simulation and abstraction in metaphor processing. *Memory & Cognition*, 49(3), 557–570. doi:10.3758/s13421-020-01109-2

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(4), 354–361. doi:10.1177/073428299901700405

Baggio, G. (2018). *Meaning in the brain*. Cambridge, MA: MIT Press. doi:10.7551/mitpress/11265.001.0001

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

Blasko, D. G., & Briihl, D. S. (1997). Reading and recall of metaphorical sentences: Effects of familiarity and context. *Metaphor and Symbol*, 12(4), 261–285. doi:10.1207/s15327868ms1204_4

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 295–308. doi:10.1037/0278-7393.19.2.295

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(11), 2669–2683. doi:10.1016/j.neuropsychologia.2012.07.021

Bowdle, B., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193–216. doi:10.1037/0033-295X.112.1.193

Cacciari, C., & Glucksberg, S. (1995). Understanding idioms: Do visual images reflect figurative meanings? *European Journal of Cognitive Psychology*, 7(3), 283–305. doi:10.1080/09541449508402450

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(2), 172–188. doi:10.1016/j.jml.2006.11.006

Chiappe, D. L., Kennedy, J. M., & Chiappe, P. (2003a). Aptness is more important than comprehensibility in preference for metaphors and similes. *Poetics*, 31(1), 51–68. doi:10.1016/S0304-422X(03)00003-2

Chiappe, D. L., Kennedy, J. M., & Smykowski, T. (2003b). Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol*, 18(2), 85–105. doi:10.1207/S15327868MS1802_2

Chiappe, D. L., & Kennedy, J. M. (1999). Aptness predicts preference for metaphors or similes, as well as recall bias. *Psychonomic Bulletin & Review*, 6(4), 668–676. doi:10.3758/BF03212977

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(6), 1445–1455. doi:10.3758/bf03193614

Christensen, R. H. B. (2019). ordinal: Regression models for ordinal data. R package version 2019. 12. https://CRAN.R-project.org/package=ordinal

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. (2002). Metaphoric interpretations of comparison-based combinations. *Metaphor and Symbol*, 17(3), 161–178. doi:10.1207/S15327868MS1703_1

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.

Gibbs, R. W., Jr., & Bogdonovich, J. (1999). Mental imagery in interpreting poetic metaphor. *Metaphor and Symbol*, 14 (1), 37–54. doi:10.1207/s15327868ms1401_4

Glicksohn, J., and Goodblatt, C. 2021 Empirical studies of poetic metaphor Kuiken, D., and Jacobs, A. M. eds. Handbook of empirical literary studies (Berlin/Boston: De Gruyter) 121–143 doi:10.1515/9783110645958-006

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, 97(1), 3–18. doi:10.1037/0033-295X.97.1.3

Glucksberg, S., & McGlone, M. S. (1999). When love is not a journey: What metaphors mean. *Journal of Pragmatics*, 31 (12), 1541–1558. doi:10.1016/S0378-2166(99)00003-X

Gray, M. E., & Holyoak, K. J. (2020). Individual differences in relational reasoning. *Memory & Cognition*, 48(1), 96–110. doi:10.3758/s13421-019-00964-y

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin*, 144(6), 641–671. doi:10.1037/bul0000145

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. (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.

Jacobs, A. M., & Kinder, A. (2018). What makes a metaphor literary? Answers from two computational studies. *Metaphor and Symbol*, 33(2), 85–100. doi:10.1080/10926488.2018.1434943

Jamrozik, A., McQuire, M., Cardillo, E. R., & Chatterjee, A. (2016). Metaphor: Bridging embodiment to abstraction. *Psychonomic Bulletin & Review*, 23(4), 1080–1089. doi:10.3758/s13423-015-0861-0

Jones, L. L., & Estes, Z. (2005). Metaphor comprehension as attributive categorization. *Journal of Memory and Language*, 53(1), 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(1), 18–32. doi:10.1016/j.jml.2006.02.004

Katz, A. N., Paivio, A., Marschark, M., & Clark, J. M. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity*, 3(4), 191–214. doi:10.1207/s15327868ms0304_1

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

Kintsch, W., & Mangalath, P. (2011). The construction of meaning. *Topics in Cognitive Science*, 3(2), 346–370. doi:10.1111/j.1756-8765.2010.01107.x

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, 7(2), 257–266. doi:10.3758/BF03212981

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

Liu, S., Erkkinen, M. G., Healey, M. L., Xu, Y., Swett, K. E., Chow, H. M., & Braun, A. R. (2015). Brain activity and connectivity during poetry composition: Toward a multidimensional model of the creative process. *Human Brain Mapping*, 36(9), 3351–3372. doi:10.1002/hbm.22849

Marks, L. E. (1982). Synesthetic perception and poetic metaphor. *Journal of Experimental Psychology: Human Perception and Performance*, 8(1), 15–23. doi:10.1037/0096-1523.8.1.15

Miljković, B. (1956/2005). *Uzalud je budim [I wake her in vain]*. Belgrade, Serbia: Narodna knjiga.

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

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: Routledge. doi:10.4324/9781351060912-4

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), 1–19. doi:10.1080/10926480802568422

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/

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(1), 600–610. doi:10.1016/j.neuroimage.2012.06.022

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(3), 800–812. doi:10.3758/s13428-014-0502-y

Sana, F., Park, J., Gagné, C. L., & Spalding, T. L. (2021). The interplay between inhibitory control and metaphor conventionality. *Memory & Cognition*, 49(6) Advance online publication. 1267–1284. 10.3758/s13421-021-01152-7

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

Stamenković, D., Ichien, N., & Holyoak, K. J. (2020). Individual differences in comprehension of contextualized metaphors. *Metaphor and Symbol*, 35(4), 285–301. doi:10.1080/10926488.2020.1821203

Stamenković, D., Milenković, K., & Dinčić, J. (2019b). Studija normiranja književnih i neknjiževnih metafora iz srpskog jezika [A norming study of Serbian literary and nonliterary metaphors]. *Zbornik Matice srpske za filologiju i lingvistiku*, 62(2), 89–104.

Tourangeau, R., & Rips, L. (1991). Interpreting and evaluating metaphors. *Journal of Memory and Language*, 30(4), 452–472. doi:10.1016/0749-596X(91)90016-D

Tourangeau, R., & Sternberg, R. J. (1981). Aptness in metaphor. *Cognitive Psychology*, 13(1), 27–55. doi:10.1016/0010-0285(81)90003-7

Utsumi, A. (2007). Interpretive diversity explains metaphor–simile distinction. *Metaphor and Symbol*, 22(4), 291–312. doi:10.1080/10926480701528071

Vartanian, O. (2012). Dissociable neural systems for analogy and metaphor: Implications for the neuroscience of creativity. *British Journal of Psychology*, 103(3), 302–316. doi:10.1111/j.2044-8295.2011.02073.x

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

Wong, E. F., & Holyoak, K. J. (2021). Cognitive and motivational factors driving sharing of internet memes. *Memory & Cognition*, 49(5) Advance online publication. 863–872. 10.3758/s13421-020-01134-1