Advancing with Age: Older Adults Excel in Comprehension of Novel Metaphors

IN PRESS – Psychology & Aging


Nicholas Ichien
nichien@sas.upenn.edu

Dušan Stamenković
dusan.stamenkovic@sh.se

Mary C. Whatley
mcwhatley@g.ucla.edu

Alan D. Castel
castel@ucla.edu

Keith J. Holyoak
kholyoak@g.ucla.edu

1Department of Psychology, University of Pennsylvania
2School of Culture and Education, Södertörn University, Stockholm
3Department of Psychology, University of California, Los Angeles

Author note:
The work was supported by National Science Foundation grant BCS-2022369 and a gift from Canon Corporation to K.J.H. We thank Miranda Calacsan for her help with experimental task design. The ideas and data appearing in the manuscript have been disseminated as a preprint: https://doi.org/10.31234/osf.io/ry9js.
The experimental stimuli, data, and analysis code for both studies are publicly accessible at: https://osf.io/je5ay/. Studies 1 and 2 were preregistered at AsPredicted #127579 (https://aspredicted.org/mz5q7.pdf) and #134910 (https://aspredicted.org/gs4m3.pdf), respectively.

Author Contributions
Conceptualization: AC, KJH; Data curation: NI; Formal analysis: NI; Funding acquisition: KJH; Investigation: NI, DS, MW; Methodology: NI, DS, MW, KJH, AC; Project administration: NI, MW; Resources: NI, DS; Software: NI; Supervision: KJH, AC; Validation: NI; Visualization: NI; Writing—original draft: NI, DS, MW, AC, KJH; Writing—review and editing: NI, DS, MW, AC, KJH.

Corresponding author:
Nicholas Ichien
Department of Psychology
University of Pennsylvania
3720 Walnut St.
Philadelphia, PA 19104
nichien@sas.upenn.edu
Abstract

Older adults may experience certain forms of cognitive decline, but some forms of semantic memory remain intact in older age. To address how metaphor comprehension changes with age, and whether metaphor comprehension relies more heavily on analogical reasoning (supported by fluid intelligence) or on conceptual combination (supported by crystalized intelligence), we compared performance of younger and older adults. In two experiments, healthy older adults (54–88 years) scored lower on a measure of fluid intelligence (Ravens Progressive Matrices) but higher on a measure of crystalized intelligence (Mill Hill Vocabulary Test) relative to younger adults (18–34 years). Groups were equally successful in comprehending relatively easy metaphors (Study 1), but older adults showed a striking advantage over younger adults for novel literary metaphors (Study 2). Mixed-effects modeling showed that measures of fluid and crystalized intelligence each made separable contributions to metaphor comprehension for both groups, but older adults relied more on crystalized intelligence than did younger adults. These age-related dissociations clarify cognitive effects of aging, and highlight the importance of crystalized intelligence for metaphor comprehension in both younger and older adults.

Keywords: aging; metaphor; fluid intelligence; crystalized intelligence
Public Significance Statement

We showed that (in spite of their diminished reasoning ability) older adults are equal or superior to younger adults on tests of metaphor comprehension. The advantage of the older adults was linked to a greater reliance on their superior verbal knowledge. Our study has the encouraging implication that the ability to read and appreciate literature, including poetic metaphors, continues to develop throughout adulthood.
**Introduction**

Older adults often experience changes in memory and associated cognitive functions, with episodic memory and working memory showing more pronounced declines compared to semantic memory, which may remain intact in older age (Park & Festini, 2019; Salthouse, 2019). For example, younger and older adults rely on various forms of memory, such as working memory and semantic memory, when reading and grasping narratives, but age-related declines in narrative processing are primarily driven by declines in working memory capacity (e.g., Morrow, Leirer, & Altieri, 1992; Tun, 1989). Still, both younger and older adults may engage in efficient forms of narrative comprehension via reliance on semantic memory, through the use of general schemas or gist-based information (Artuso & Belacchi, 2021; Gallo, Hargis, & Castel, 2019; Hess, 2005; Radvansky & Dijkstra, 2007). Thus, there may be circumstances in which older adults—who often have richer vocabulary relative to younger adults and can accurately retrieve word meanings from semantic memory (Burke & Shafto, 2011; Murphy & Castel, 2021)—can utilize relevant schemas and domain knowledge to offset declines in working memory and efficiently interpret narratives and written text (cf. Jeong & Kim, 2009).

Metaphor comprehension is one form of text or narrative processing, which may rely differentially on domain-general cognitive abilities (e.g., working memory, inhibition) and semantic or verbal knowledge, suggesting potential age-related differences. In his poem “Democracy”, Langston Hughes characterizes his longing for freedom by a striking metaphor: “Freedom / is a strong seed / planted / in a great need.” As a psychological process, the comprehension of metaphor—the most prominent form of figurative speech—lies at the interface between neurally dissociable forms of processing: domain-general thought and modality-specific language comprehension (e.g., Blank & Federenko, 2017). Theories of metaphor comprehension
have tended toward two camps, each emphasizing one side of this divide (for a review see Holyoak & Stamenković, 2018). The classical view inherited from Aristotle and championed by several modern psychologists (e.g., Tourangeau & Sternberg, 1981; Gentner & Clement, 1988) is that metaphors are interpreted by analogical reasoning. Under this view, a process of structural alignment serves to map *freedom* (the target concept being described in the example from Hughes) to a *strong seed* (the more concrete source used to interpret the target) as corresponding elements in two larger systems of concepts (with *freedom* belonging to a system having to do with states of human existence and *strong seed* belonging to a system having to do with plant life). Metaphor comprehension, though it typically involves some form of language processing, thus depends on a domain-general variety of thought.

An alternative view is that metaphors are interpreted by a process of conceptual combination (e.g., Estes & Glucksberg, 2000; Keane & Costello, 2001; Kintsch, 2000). In the Hughes example, this process of conceptual combination might involve the revision of or emphasis on certain features of the concept *freedom* that would make it seem more similar to *strong seed* (perhaps emphasizing that *freedom* constitutes the necessary beginning of some new life). One variant of this view (Glucksburg & Keysar, 1990) treats this process as a kind of categorization (e.g., *seed* is interpreted not as the fertilized ovule of a flowering plant, but as a more abstract category of “source of development or growth”). Under the conceptual combination view, metaphor comprehension is less dependent on domain-general thought, and is instead an extension of the processes that support comprehension of literal language. Finally, a compromise position is that novel metaphors are interpreted using analogical reasoning, whereas familiar metaphors are understood by some form of conceptual combination (Bowdle & Gentner, 2005). In the current paper, we compare how individual differences in cognitive abilities among older and younger
adults predict metaphor comprehension ability in order to disambiguate between the views described above. We ultimately provide support for the view that metaphor comprehension is continuous with literal language comprehension, rather than relying heavily on domain-general thought.

**Individual Differences and Metaphor Comprehension**

One general approach to investigating the mechanisms of metaphor comprehension is to focus on individual differences in intellectual abilities (Cattell, 1971). Analogical reasoning is known to depend on *fluid* intelligence: the ability to manipulate complex information in working memory while inhibiting salient but task-irrelevant information (Gray & Holyoak, 2020). At the neural level, analogy and other forms of reasoning that require fluid intelligence depend on the operation of a frontoparietal network (Duncan, Assem, & Shashidhara, 2020). In contrast, conceptual combination is linked to a separate language network that includes the left anterior temporal lobe (Parrish & Pylkkänen, 2022). Verbal knowledge, which includes word meanings and the ability to combine them, constitutes *crystalized* intelligence, which broadly reflects any knowledge accumulated from past experience.

Studies of metaphor comprehension have shown that crystalized intelligence impacts comprehension of a wide range of metaphors (both literary and more conventional). In contrast, fluid intelligence tends to play a lesser role generally, though it does support metaphor comprehension for relatively challenging literary metaphors (Stamenković, Ichien, & Holyoak, 2019a, 2020; Stamenković, Milenković, Ichien, & Holyoak, 2023). This evidence suggests that conceptual combination may be the dominant process underlying metaphor comprehension, although analogy may contribute to understanding difficult metaphors.
Metaphor Comprehension Across the Lifespan

Examining metaphor comprehension across the lifespan may reveal the impact of a dissociation between fluid and crystalized intelligence, in addition to potential age-related differences in metaphor comprehension. There is considerable evidence that healthy aging is accompanied by diminished fluid intelligence (Parkin & Java, 1999; Staff, Hogan, & Whalley, 2014), which results in impairment on tests of analogical reasoning (Viskontas et al., 2004). In contrast, crystalized verbal intelligence is largely spared, or perhaps even enhanced, over the course of typical aging (Horn, 1982; Horn & Cattell, 1967; Unamath & Marsh, 2014). This dissociation appears to reflect the greater impact of aging on frontal brain areas associated with fluid intelligence than on temporal regions associated with language and crystalized intelligence (Martin, Williams, Saur, & Hartwigsen, 2023; Staff et al., 2014). Healthy older adults typically have better vocabulary than younger adults (Ben-David, Erel, Goy, & Schneider, 2015), which could facilitate some forms of metaphor comprehension. It follows that if metaphor comprehension primarily relies on analogical reasoning, this ability should decline in older age. In contrast, if metaphors are understood using some form of conceptual combination, performance should be spared or even enhanced in healthy older adults.

The evidence concerning the impact of aging on metaphor comprehension is decidedly mixed, with some studies showing evidence of impairment in processing figurative language (e.g., Uekermann, Thoma, & Daum, 2008) and others (e.g., Light, Owens, Mahoney, & La Voie, 1993; Newsome & Glucksberg, 2002) showing stable performance across ages (for a review see Bartczak, 2017). Previous findings concerning metaphor comprehension across the lifespan are difficult to interpret because of variations in age and health characteristics of the populations from
which samples were drawn, the types of metaphors or other figurative language that were studied, and the nature of the tasks administered.

The present study was designed to allow more rigorous analyses of the factors that impact metaphor comprehension and of how these factors relate to metaphor comprehension in younger and older adults. In two experiments using relatively large general-population samples of healthy younger and older adults with basic computer literacy, we obtained measures of individual differences in both fluid and crystalized intelligence. All participants were asked to interpret sets of metaphors that have been normed on several dimensions known to impact metaphor comprehension. For both younger and older participants, we examined the relations between each of the two cognitive factors and metaphor comprehension. We predicted that metaphor comprehension in both younger and older adults would rely more heavily on crystallized intelligence than on fluid intelligence (Stamenkovic et al., 2019, 2020, & 2023). We also examined the possible dissociation between fluid and crystalized intelligence with aging, together with the relative abilities of younger and older adults to grasp a range of metaphors. Based on previous studies, we expected that fluid intelligence would show age-related declines and that older adults would perform worse than younger adults on our measure of fluid intelligence (Parkin & Java, 1999; Staff, Hogan, & Whalley, 2014). On the other hand, we expected crystallized intelligence to be preserved in older adults and that older adults would match or perform better than younger adults on our measure of verbal knowledge (Horn, 1982; Horn & Cattell, 1967; Unamath & Marsh, 2014). Correspondingly, we predicted that older adults would match or exceed younger adults in comprehension of metaphors. Because poetry is a natural source of original metaphors (Lakoff & Turner, 1989; Holyoak, 2019), our stimuli included a set of poetic metaphors chosen to be novel to all participants.
Study 1

Method

Transparency and Openness

Data collection plan and analyses for Study 1 were preregistered on AsPredicted (AsPredicted #127579 https://aspredicted.org/mz5q7.pdf). Analytic code, study materials, and de-identified data on which Study 1 conclusions are based are available at the following link: https://osf.io/je5ay/.

Participants

Participants were 78 older adults ($M_{age} = 69.47$ years, $SD_{age} = 4.86$, range = [54,88]; 45 female, 31 male, 2 gender not reported) and 106 younger adults ($M_{age} = 22.80$ years, $SD_{age} = 1.87$, range = [18,26]; 44 female, 60 male, 2 nonbinary). Participants did not report their race. Sample sizes were selected to be comparable to those used in previous studies of individual differences in metaphor comprehension using young adults. All participants were recruited online via Prolific Academic, a general population very likely to be cognitively healthy and computer literate, and to provide high-quality data (Peer, Rothschild, Gordon, Evernden, & Damer, 2022). Data collection took place during the summer of 2023. Participants completed the three experimental tasks detailed below in return for payment of $8. The study was approved by the Institutional Review Board at the University of California, Los Angeles (UCLA) (IRB#18-000767 "Achieving Analogical Reasoning Via Human and Machine Learning"). We eliminated 2 older adults (final $n_{older} = 76$) and 15 younger adults (final $n_{younger} = 91$) for failing to achieve at least chance performance across all experimental tasks, or to provide a sensible response to a question asking participants to name their favorite book.
Design, Materials, and Procedure

All participants completed three tasks in a fixed order. The first two tasks assessed individual differences in cognitive abilities, and the final task involved metaphor comprehension. For convenience and to avoid asking participants to endure an overly-demanding battery of tasks, we measured fluid and crystallized with single measures: Raven’s Advanced Progressive Matrices and the Mill Hill Vocabulary test, respectively. Adopting this single-measure approach, we have previously demonstrated dissociations between fluid and crystallized intelligence with respect to their association with metaphor comprehension performance (Stamenkovic et al., 2019, 2020, 2023). We mention support for the validity of each of these measures below.

Task 1: Raven’s Progressive Matrices (RPM). A shortened, 12-item version of Raven’s Advanced Progressive Matrices (Arthur, Tubre, Paul, & Sanchez-Ku, 1999) was administered to assess fluid intelligence. Each trial of this task presents a 3x3 array of simple geometric figures that instantiate some pattern across rows and columns (e.g., progression in number of figure components from the leftmost to middle and to the rightmost column) but that omits the bottom-right cell of that array, and participants are tasked with selecting from among 8 options the figure that best fills that bottom-right cell. The RPM is generally considered a central measure of fluid intelligence (Snow, Kyllonen, & Marshalek, 1984), and scores on this test correlate with performance on a suite of other relational reasoning tasks (Gray & Holyoak, 2020). The degree to which performance on the RPM predicts individual differences in metaphor comprehension should thus reflect the extent to which explicit analogical reasoning is required to comprehend metaphors.

Task 2: Mill Hill Vocabulary Scale (MH). A shortened 20-item version of the Mill Hill Vocabulary Scale (Raven, 1965) was administered to assess crystallized verbal knowledge. On each trial of this vocabulary task, participants are shown a target word and are asked to select from
a set of 6 options the one that constitutes the closest synonym to the target word. This abbreviated task correlates strongly with performance on another standard measure of verbal knowledge, WAIS-III (Wechsler, 1997), and performance on this abbreviated task has been used to compare verbal knowledge across older and younger age groups (Ben-David et al., 2015). The extent to which performance on the MH predicts individual differences in metaphor comprehension should therefore reflect how much metaphor comprehension relies on verbal knowledge.

Task 3: Metaphor comprehension. The third task assessed metaphor comprehension using 15 literary metaphors selected from a list of literary metaphors drawn from poetry anthologies by Katz et al. (1988), and 15 non-literary metaphors adapted from four-term $A:B::C:D$ analogy problems developed by Green et al. (2010, 2012) (e.g., childhood : life :: morning : day). Table 1 provides examples of each kind of metaphor problem. The selected metaphors instantiated some variant of a nominal syntactic form ($X$ is $Y$; e.g., with an adjective modifier or with a prepositional phrase), and were rated high on a goodness scale (Stamenković et al., 2019a). These were the same materials used in Stamenković et al. (2019a) Study 2a and Study 2b, which showed that in younger adults, comprehension of literary metaphors relied on both fluid intelligence and crystalized verbal knowledge, whereas comprehension of non-literary metaphors only relied on the latter.
Table 1. Example items from metaphor comprehension task (Study 1). Participants were asked to select the best interpretation of each metaphor from among a set of three (the correct answer and two distractors).

<table>
<thead>
<tr>
<th>Type</th>
<th>Metaphor</th>
<th>Correct</th>
<th>Distractor 1</th>
<th>Distractor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literary</td>
<td>Water is the blood of soft snows.</td>
<td>Water originates from soft snows.</td>
<td>Water brings coldness.</td>
<td>Soft snows are thicker than water.</td>
</tr>
<tr>
<td></td>
<td>Man is a leaf in the gardens of God.</td>
<td>God cherishes human beings.</td>
<td>God waters the soil.</td>
<td>Human beings love God.</td>
</tr>
<tr>
<td></td>
<td>The soul is a rope that binds heaven and earth.</td>
<td>The soul allows one to travel from earth to heaven.</td>
<td>The soul contains both heaven and earth.</td>
<td>The soul is what makes heaven look like earth.</td>
</tr>
<tr>
<td>Non-literate</td>
<td>Childhood is the morning of life.</td>
<td>Childhood comes early in life.</td>
<td>Childhood is initiated before life.</td>
<td>Childhood comes at the same time as life.</td>
</tr>
<tr>
<td></td>
<td>A wire is the riverbed of electricity.</td>
<td>Electricity goes through wires.</td>
<td>Electricity is the source of wires.</td>
<td>Electricity goes through riverbeds.</td>
</tr>
<tr>
<td></td>
<td>Invention is the child of an inventor.</td>
<td>Inventors are the creators of inventions.</td>
<td>Inventors are like children.</td>
<td>Inventors neglect inventions.</td>
</tr>
</tbody>
</table>

On each trial of this task, participants were shown a metaphor and three potential interpretations of that metaphor, and they were instructed to “choose the interpretation that is closest in meaning to the metaphor” (see Table 1). We presented literary and non-literary metaphors in two separate blocks to avoid a kind of strategy spillover (e.g., participants might adopt an analogy-based approach to interpret literary metaphors but then redundantly applying it to nonliterary metaphors presented in the same block). Blocks were presented in a random order for each participant (i.e., literary followed by non-literary or vice versa) and items were displayed in a randomized order within each block. The distractors in the metaphor comprehension task were created to be potentially relevant but misleading interpretations of the metaphors, aiming to challenge the participant’s ability to discern the intended figurative meaning. They vary in plausibility and semantic distance from the correct answer, and they consist both of literal interpretations of the metaphor (e.g., “electricity goes through riverbeds” is a distractor that
interprets the metaphor “a wire is the riverbed of electricity” literally) and of incorrect figurative alternatives (e.g., “electricity is the source of wires” is a poor attempt at a figurative interpretation of “a wire is the riverbed of electricity”).

**Results**

Figure 1 depicts performance on the measures of fluid intelligence (RPM) and of crystalized intelligence (MH) for the two age groups. Consistent with the pattern observed in past research (Parkin & Java, 1999; Staff, Hogan, & Whalley, 2014), a Wilcoxon rank-sum test showed that younger adults scored higher ($M_{younger} = 0.52$, $SD_{younger} = 0.21$) than did older adults on the RPM task measuring fluid intelligence ($M_{older} = 0.43$, $SD_{older} = 0.18$; $W = 2672$, $p = .011$). In contrast, verbal knowledge was not only spared but enhanced in the older group: older adults outperformed ($M_{older} = 0.77$, $SD_{older} = 0.14$) younger adults on the MH vocabulary assessment ($M_{younger} = 0.63$, $SD_{younger} = 0.14$; $W = 5264.5$, $p > .001$) (Horn, 1982; Horn & Cattell, 1967). Mill Hill performance converges with that reported in Ben-David et al. (2015), who used community samples of older adults ($n_{older} = 737$, $M_{older} = 0.75$, $SD_{older} = 0.10$) and younger adults ($n_{younger} = 1299$, $M_{younger} = 0.66$, $SD_{younger} = 0.10$).
Figure 1. Proportion correct on tests of fluid intelligence (RPM; left panel) and verbal knowledge (MH; right panel), broken down by age group (Study 1). Individual points reflect performance of individual participants, horizontal lines reflect mean proportions, and boxes reflect 95% confidence intervals. Plots were generated using Mika Braginsky’s `ggpirate` R package.

Figure 2 (left panel) summarizes performance on the test of metaphor comprehension. We analyzed metaphor comprehension by fitting a logistic mixed-effects model to trial-level accuracy, using the `glmer` function from version 1.1.26 of the LME4 R package (Bates et al., 2015) in R version 4.3.1 (R. Core Team, 2021). Unlike conventional logistic regression, the mixed-effects modeling that we employed enabled us to statistically control for variability in metaphor comprehension attributable to idiosyncrasies in individual participants, as well as of individual metaphor comprehension items. Compared to conventional logistic regression, such modeling
provides more assurance that conclusions from statistical tests generalize beyond our particular sample and beyond the particular items used (Barr, Levy, Scheepers, & Tily, 2013; Clark, 1973).

We defined a full model including three-way interaction terms for metaphor type (literary vs. non-literary) * age group (older vs. younger) * fluid intelligence (RPM score) and for metaphor type * age group * verbal knowledge (MH score) and all lower-level terms (e.g., metaphor type, age group, fluid intelligence, metaphor type * age group, metaphor type * fluid intelligence). This model features a ‘maximal’ random effect structure, given our experimental design (Barr et al., 2013), consisting of (1 + metaphor type | participant) and (1 + age group | metaphor problem) as random-effect terms. Note that the difference between the present logistic mixed-effects model and a conventional logistic regression model is the inclusion of those random effect terms.

We used likelihood-ratio tests to compare this full model to each of two reduced models that respectively omitted the fluid intelligence and verbal knowledge three-way interaction terms but that were otherwise equivalent to the full model. These comparisons showed that dropping either term did not significantly increase model prediction error (fluid intelligence \( \Delta AIC = 2.0, \chi^2 (1) = .02, p = .875 \); verbal knowledge: \( \Delta AIC = 0.9, \chi^2 (1) = 1.13, p = .288 \)). Still, in order to test predicted simple effects that compare metaphor comprehension across age groups within each metaphor type and that compare performance across metaphor types within each age group, we estimated marginal trends from the full model described above (with interaction terms).
Figure 2. Metaphor comprehension proportion correct (Study 1), broken down according to metaphor type and age group (top), and by performance on individual difference measures of fluid intelligence (yellow) and verbal knowledge (blue), separated by metaphor type and age group (bottom). On both plots, individual points reflect performance of individual participants, and on
the left plot, horizontal lines reflect mean proportions and boxes reflect 95% confidence intervals. This plot was generated using Mika Braginsky’s ggpirate R package.

We compared estimated marginal means of the full model using the emmeans and pairs functions from version 1.8.4 of the emmeans R package (Lenth, 2023). The left panel of Figure 2 presents accuracy rates on the metaphor comprehension task, broken down according to metaphor type and age group. For both literary and non-literary metaphors, there were no performance differences across age groups (literary: $M_{\text{older}} = .75$, $SD_{\text{older}} = .13$; $M_{\text{younger}} = .74$, $SD_{\text{younger}} = .12$; difference estimate = .07, SE = .18, $z = .41$, $p = .68$; non-literary: $M_{\text{older}} = .92$, $SD_{\text{older}} = .10$; $M_{\text{younger}} = .92$, $SD_{\text{younger}} = .10$; difference estimate = .27, SE = .28, $z = .98$, $p = .32$). For both older and younger adults, literary metaphors were more difficult than non-literary metaphors (older: difference estimate = 1.56, SE = .38, $z = 4.14$, $p < .001$; younger: difference estimate = 1.75, SE = .32, $z = 5.45$, $p < .001$). Thus, although there were no performance differences across age groups, both age groups found the literary metaphors more difficult, replicating the findings of Stamenković et al. (2019a) and extending them to older adults.

We then tested simple trends reflecting the extent to which metaphor comprehension relied on each of the two individual difference measures, using the emtrends and test functions from the emmeans R package (Lenth, 2023). The right panel of Figure 2 plots participant-level metaphor comprehension accuracy against their fluid intelligence (RPM performance) and verbal knowledge (MH performance). Metaphor comprehension did not rely on fluid intelligence within any combination of age group and metaphor type (older-literary: estimated trend = 0.30, SE = 0.53, $z = 0.56$, $p = .574$; older-non-literary: estimated trend = 0.87, SE = 0.96, $z = 0.91$, $p = .366$; younger-literary: estimated trend = 0.58, SE = 0.38, $z = 1.55$, $p = .122$; younger-non-literary: estimated trend = 0.96, SE = 0.69, $z = 1.39$, $p = .164$), but it did rely on verbal knowledge across all combinations of age group and metaphor type (older-literary: estimated trend = 2.25, SE = 0.69, $z = 3.27$, $p =$
This result replicates the finding of Stamenković et al. (2019a) that verbal knowledge predicts metaphor comprehension in younger adults and extends it to older adults. As in the earlier study, verbal knowledge was the dominant predictor overall. Stamenković et al. found that fluid intelligence was an additional reliable predictor for younger adults when tested with literary metaphors; in contrast, the present study did not show any reliable influence of fluid intelligence on younger adults’ comprehension of literary metaphors. The two studies involved different populations: Stamenković et al. drew a sample from college undergraduates, whereas the present study sampled the general population via Prolific Academic. It is possible that college students are more likely than the general population to make use of an effortful analogy-based strategy.

Finally, in an exploratory analysis, we assessed whether age *per se* had any impact on metaphor comprehension for older adults, over and above individual differences in fluid and crystalized intelligence. To do so, we fit a logistic mixed-effects model to older adults’ metaphor comprehension performance with *chronological age* \( (M = 69.47 \text{ years}, SD = 4.86, \text{range} = [54,88]) \) as a fixed effect of interest, and with *fluid intelligence * metaphor type and *verbal knowledge * metaphor type interaction terms (along with lower-level terms), and featuring a maximal random-effect structure (Barr et al., 2013), with \((1 + \text{metaphor type} | \text{participant})\) and \((1 | \text{metaphor problem})\) random-effect terms. (Note that a \((1 + \text{age group} | \text{metaphor problem})\) random effect was not possible as the present analysis was restricted to older adults.) *Chronological age* was not a reliable predictor of metaphor comprehension in this model (estimated trend = -0.01, \(SE = 0.02, z\)
Thus, after controlling other variables, success in metaphor comprehension was independent of differences in chronological age among older adults.

**Study 2**

In Study 1, metaphor comprehension was very similar across both younger and older adults, with equal comprehension scores and comparable sensitivity to individual differences in verbal knowledge. However, the lack of differences between the age groups may reflect the fact that the metaphors were relatively easy (about 75% accuracy even for the more difficult literary metaphors). In addition, the sample size in Study 1 was chosen to be sensitive to individual differences within an age group, and may have been too small to detect any difference between the patterns of individual differences across age groups (e.g., whether or not the extent to which crystalized intelligence predicted metaphor comprehension in older adults differed from that in younger adults).

To address the limitations of Study 1, Study 2 introduced a more difficult metaphor comprehension task using translations of unfamiliar metaphors drawn from Serbian poetry. In addition, we doubled the sample sizes to increase statistical power.

**Method**

_Transparency and Openness_

Data collection plan and analyses for Study 2 were preregistered on AsPredicted (AsPredicted #134910 [https://aspredicted.org/gs4m3.pdf](https://aspredicted.org/gs4m3.pdf)). Analytic code, study materials, and de-identified data on which Study 2 conclusions are based are available at the following link: https://osf.io/je5ay/ (Ichien, 2024).
Participants

Participants were 156 older adults (ages: $M_{\text{older}} = 69.48$ years, $SD_{\text{older}} = 4.11$, $\text{range}_{\text{older}} = [64, 85]$; 1 nonbinary, 89 female, 64 male, 3 gender not reported) and 161 younger adults (ages: $M_{\text{younger}} = 22.48$ years, $SD_{\text{younger}} = 2.06$, $\text{range}_{\text{younger}} = [18, 34]$; 85 female, 1 male; 4 nonbinary). Participants did not report their race. As in Study 1, all participants were recruited online via Prolific Academic, and they completed the three experimental tasks detailed below in return for payment of $8. Data collection took place during the summer of 2023. The study was approved by the Institutional Review Board at UCLA (IRB#18-000767 "Achieving Analogical Reasoning Via Human and Machine Learning"). We eliminated 10 older adults (final $n_{\text{older}} = 146$) and 19 younger adults (final $n_{\text{younger}} = 142$) for failing to achieve at least chance performance across all experimental tasks, or to provide a sensible response to a question asking participants to name their favorite book.

Design, Materials, and Procedure

All participants completed three tasks in a fixed order. The first two tasks were the same as in Study 1. The third task involved the same procedure, but the selection of metaphors involved more difficult expressions.

**Task 3: Poetic metaphor comprehension (English translations of metaphors drawn from Serbian poetry).** The third and final task involved a set of 30 Serbian literary metaphors translated into English. These were selected from a list of literary metaphors drawn from Serbian poetry and normed on several dimensions/features (Stamenković et al., 2019b), including aptness, familiarity, perceived quality, and metaphoricity. The norming study was based on 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, Vojisav 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. All metaphors were then converted to the *<nominal> is <nominal>* form. In a subsequent study (Milenković et al., 2024), these 55 metaphors were translated into English by two translators, with a third translator verifying the translations. The translated metaphors were again normed (by different participants) on four key aspects: metaphor quality, aptness, metaphoricity, and familiarity. Given the known importance of aptness and familiarity for metaphor comprehension (see Stamenković et al., 2023), we used these two dimensions to select 30 (out of 55) items for the present study. High aptness and low familiarity were combined and weighted equally, and the 30 highest-ranked metaphors were selected. Our aim was to generate a set of difficult but highly apt metaphors that were generally unknown to English speakers. As in Study 1, on each trial of this task, participants were shown a metaphor and three potential interpretations of that metaphor, and they were instructed to “choose the interpretation that is closest in meaning to the metaphor” (see Table 2). The items were displayed in a randomized order for each participant.

**Table 2.** Examples of comprehension task with literary metaphors (Study 2). Participants were asked to select the best interpretation of each metaphor from among a set of three (the correct answer and two distractors).

<table>
<thead>
<tr>
<th>Metaphor</th>
<th>Correct</th>
<th>Distractor 1</th>
<th>Distractor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood is powerless ink.</td>
<td>Sacrifice achieves little in comparison to the impact of words and ideas.</td>
<td>Writing symbolizes an activity that holds the potential to transform into violence.</td>
<td>The potency of weapons often results in senseless bloodshed.</td>
</tr>
<tr>
<td>The deceased is a star above a void.</td>
<td>Death is an escape from the abyss of life.</td>
<td>The deceased become stars in the sky.</td>
<td>People who have passed away return to nature in a new form.</td>
</tr>
<tr>
<td>The soul is a narrow dungeon.</td>
<td>The soul limits one’s true potential and emotions.</td>
<td>The soul is very narrow, and we can get trapped.</td>
<td>People are often trapped by negative experiences.</td>
</tr>
</tbody>
</table>
**Results**

Figure 3 depicts performance on the measures of individual differences for each age group. As in Study 1, younger adults outperformed \( (M_{younger} = 0.54, SD_{younger} = 0.22) \) older adults on the RPM \( (M_{older} = 0.42, SD_{older} = 0.18; W = 7233.5, p > .001) \), whereas older adults \( (M_{older} = 0.78, SD_{older} = 0.10) \) outperformed younger adults on the MH \( (M_{younger} = 0.66, SD_{younger} = 0.14; W = 16084, p > .001) \). Besides replicating the pattern obtained in Study 1, this finding is consistent with other previous studies (Parkin & Java, 1999; Staff, Hogan, & Whalley, 2014, Horn, 1982; Horn & Cattell, 1967). As in Study 1, Mill Hill performance converges with that exhibited by community samples of older adults \( (n_{older} = 737, M_{older} = 0.75, SD_{older} = 0.10) \) and younger adults \( (n_{younger} = 1299, M_{younger} = 0.66, SD_{younger} = 0.10) \) (Ben-David et al., 2015).

*Figure 3. Proportion correct on tests of fluid intelligence (RPM; left panel) and verbal knowledge (MH; right panel), broken down by age group (Study 2). Individual points reflect performance of*
individual participants, horizontal lines reflect mean proportions, and boxes reflect 95% confidence intervals. Plots were generated using Mika Braginsky’s ggpirate R package.

Figure 4 (left panel) summarizes performance on the test of metaphor comprehension. In order to analyze metaphor comprehension, we fit a logistic mixed-effects model to performance on individual trials on the metaphor comprehension task, using the glmer function from version 1.1.26 of the LME4 R package (Bates et al., 2015). We defined a full model including two-way interaction terms for age group (older vs. younger) * fluid intelligence (RPM score) and for age group * verbal knowledge (MH score) and all lower-level terms (age group, fluid intelligence, etc.), and featuring the ‘maximal’ random-effect structure, given our experiment design (Barr et al., 2013), consisting of (1 | participant) and (1 + age group | metaphor problem) random effect terms. We used a likelihood-ratio test to compare this full model to an otherwise equivalent model omitting the age group term. As shown in the left panel of Figure 4, metaphor comprehension was higher for older adults (Molder = 70.42, SDolder = .11) than for younger adults (Myounger = 64.52, SDyounger = .11; ΔAIC = 6.40, χ² (3) = 12.39, p = .006). Strikingly, the more difficult metaphors and increased statistical power in Study 2 enabled us to detect an advantage for older adults compared to younger adults in metaphor comprehension ability.

We then used this same approach to compare the full model to each of two reduced models that respectively omitted the fluid intelligence or verbal knowledge interaction terms but that were otherwise equivalent to the full model. These comparisons showed that dropping the fluid intelligence interaction term did not reduce model prediction error, ΔAIC = 1.90, χ² (1) = .12, p = .729, but that dropping the verbal knowledge interaction term did: ΔAIC = 3.70, χ² (1) = 5.28, p = .022). These results indicate that whereas older and younger adults did not reliably differ in the extent to which fluid intelligence predicted metaphor comprehension, they did differ in the extent to which verbal knowledge predicted metaphor comprehension.
Figure 4. Metaphor proportion correct (Study 2), broken down according to age group (top), and by participant-level performance on individual difference measures of fluid intelligence (yellow) and verbal knowledge (blue), separated by age group (bottom). Individual points reflect
performance of individual participants; and on the upper left plot, horizontal lines reflect mean proportions and boxes reflect 95% confidence intervals. These plots were generated using Mika Braginsky’s ggpirate R package.

We then obtained estimated marginal trends from the full model to clarify the impact of individual differences on metaphor comprehension within each age group. We tested simple trends reflecting the extent to which metaphor comprehension relied on each individual difference measure, using the emtrends and test functions from the emmeans R package (Lenth, 2023). The right panel of Figure 4 plots participant-level metaphor comprehension accuracy against their fluid intelligence (RPM performance) and verbal knowledge (MH performance). For both age groups, metaphor comprehension depended both on fluid intelligence (older: estimated trend = 1.25, SE = .27, z = 4.70, p < .001; younger: estimated trend = 1.14, SE = 0.20, z = 5.73, p < .001), and on verbal knowledge (older: estimated trend = 2.70, SE = 0.46, z = 5.91, p < .001; younger: estimated trend = 1.41, SE = 0.32, z = 4.41, p < .001). Thus, despite the decline in fluid intelligence among older adults, they still relied in part on their general reasoning ability to understand difficult metaphors, as did younger adults. And as observed in Study 1 (also Stamenković et al., 2019a), verbal knowledge played a major role in metaphor comprehension for both age groups. Notably, given the numerically larger trend in verbal knowledge for older adults than younger adults and given the significant verbal knowledge x age group interaction mentioned in the previous paragraph, the present analyses indicate that older adults relied more heavily on their verbal knowledge in comprehending metaphors than did younger adults. Not only is verbal knowledge enhanced in older adulthood, but older adults successfully lean into their verbal strength to comprehend difficult and unfamiliar metaphors, more than compensating for their decline in fluid intelligence.
Finally, as in Study 1, we carried out an exploratory analysis to assess whether age had any impact on metaphor comprehension in older adults, beyond individual differences in the cognitive abilities we measured. We fit a logistic mixed-effects model to older adults’ metaphor comprehension performance, with chronological age ($M = 69.48$ years, $SD = 4.11$, range $= [64,85]$) as a fixed effect of interest, and with fluid intelligence and verbal knowledge as covariates, and ($1 \mid \text{participant}$) and ($1 \mid \text{metaphor problem}$) random effect-terms, which constitute the maximal random-effect structure (Barr et al., 2013). As in Study 1, chronological age was not a reliable predictor of metaphor comprehension in this model (estimated trend $= -.01$, $SE = .01$, $z = 1.10$, $p = .27$), again indicating that metaphor comprehension was independent of chronological age per se within our older adult sample.

**Discussion**

While there are a variety of changes and declines in cognitive function in older age, relatively healthy older adults may rely on intact or even enhanced semantic memory and knowledge. The present research examined how aging is related to metaphor comprehension and the extent to which metaphor comprehension depends on crystalized and fluid intelligence in younger and older adults. Consistent with past work, the two studies reported here demonstrate that although fluid intelligence is diminished in older adults (Parkin & Java, 1999; Staff, Hogan, & Whalley, 2014, Horn, 1982; Horn & Cattell, 1967), crystalized verbal intelligence is enhanced (Horn, 1982; Horn & Cattell, 1967; Unamath & Marsh, 2014). Study 1 found that for relatively simple metaphors, comprehension depended solely on verbal knowledge, and was comparable for younger and older adults. Study 2 examined metaphor comprehension for more complex and unfamiliar literary metaphors created by translating from Serbian poetry into English. Fluid and crystalized intelligence each made separable contributions to metaphor comprehension for both
age groups. However, for older adults the measure of crystalized intelligence (on which they excelled) made a stronger contribution to their overall superior metaphor comprehension. These findings suggest that the strength of older adults’ verbal knowledge may compensate for weaker fluid intelligence, enabling older adults to achieve higher accuracy than younger adults in understanding complex metaphors. We acknowledge that in the present research we used only a single task to measure each of fluid intelligence (Raven’s Advanced Progressive Matrices) and crystallized intelligence (the Mill Hill Vocabulary Test). We adopted this approach to avoid an overly-demanding experimental session (especially for older adults), and note that both tasks are well-validated measures of their associated construct (Ben-David et al., 2015; Gray & Holyoak, 2020; Snow et al., 1984).

**Implications for Metaphor Comprehension**

These findings add to a growing body of evidence that the comprehension of metaphors depends most fundamentally on conceptual combination—the same systematic merging of word meanings central to understanding literal language (Kintsch, 2000). Both for younger adults (Stamenković et al., 2019a, 2020, 2023), and younger and older adults in the present study, individual differences in verbal knowledge predict success in comprehension task for all types of metaphors. In contrast, fluid intelligence (on which analogical reasoning depends) comes into play primarily as a secondary predictor of performance for complex and unfamiliar metaphors.

The evidence from older adults is particularly compelling because it indicates a double dissociation between the impact of aging on analogical reasoning versus metaphor comprehension: although older adults tend to perform worse than younger adults on tests of analogical reasoning (Viskontas et al., 2004), they can comprehend complex metaphors more accurately than do younger adults. An intriguing observation is that the opposite dissociation has been shown in
studies that compare the performance of typically-developing individuals to that of age- and IQ-
matched individuals with Autism Spectrum Disorder (ASD). Meta-analyses support the
conclusion that although individuals with ASD show spared or sometimes enhanced performance
on tests of analogical reasoning (Morsanyi, Stamenković, & Holyoak, 2020a), they are impaired
on tests of metaphor comprehension (Morsanyi, Stamenković, & Holyoak, 2020b). These twin
dissociations support a common conclusion: metaphor comprehension primarily depends on the
language system (spared in normal aging, impaired in ASD), not the general reasoning system
(impaired in normal aging, spared in ASD).

The present findings help to reconcile opposing claims about age-related changes in
inhibitory control that impact metaphor processing. Fluid intelligence includes both working-
memory capacity and inhibitory control, both of which decline during typical aging (Borella et al.,
2008; Hasher & Zacks, 1988; McCabe et al., 2010; but see Rey-Mermet & Gade, 2018). When
people are required to decide whether sentences are literally true or false, older adults have more
difficulty than younger adults in reporting that those with a metaphorical interpretation are false
(Morrone, Declercq, Novella, & Besche, 2010). Yet after processing a metaphorical expression,
older as well as younger adults automatically inhibit features of the target that are not consistent
with its metaphorical interpretation (Newsome & Glucksberg, 2002). Both of these results support
the conclusion that metaphors are interpreted automatically by the language system. Inhibition of
metaphor-irrelevant features likely involves the same linguistic processes that resolve the
interpretation of ambiguous words, and does not depend on executive control. Older and younger
adults therefore display a similar tendency to inhibit irrelevant features. At the same time, because
they have a deficit in inhibitory control, older adults have difficulty suppressing a metaphorical
interpretation automatically generated by the language system.
Implications for Successful Aging

Older populations exhibit great variability in their cognitive and other abilities (Patel et al., 2022; Tucker-Drob & Salthouse, 2011), a fact that likely accounts at least in part for the variability of findings concerning metaphor processing by older adults. The aging population is highly diverse in general health, physical and mental activity levels, and lifestyles. The samples of older adults for our two studies were drawn from Prolific Academic and tested online. Our participants were able and willing to continue to seek some paid employment, and had acquired (probably in later life) the skills necessary to perform computer-based tasks (see also Greene & Naveh-Benjamin, 2022, for a discussion of this issue and sample). Thus, our samples may be considered to represent what is generally considered successful aging (Castel, 2018; Nyberg & Pudas, 2019; Urtamo et al., 2019). Broadly speaking, our participants are in a way survivors, likely with above-average capabilities relative to their age cohort. This assessment is supported by the fact that the superior metaphor comprehension of the older relative to the younger group in Study 2 was reliable even after controlling for fluid and crystalized intelligence. Future research could assess metaphor comprehension in a more diverse sample, and one that is not recruited and tested on-line.

We certainly cannot conclude that metaphor comprehension is inevitably enhanced by aging; rather, superior ability can be considered a sign of successful aging (cf. Nyberg & Pudas, 2019). It is notable that in both of the present experiments, chronological age within the older samples had no reliable relation to success on the tests of metaphor comprehension after accounting for the measured cognitive differences. Maintenance of cognitive abilities—especially verbal knowledge—is the key to preserving and enhancing the ability to deal with figurative language. If this prerequisite is met, there is reason to be optimistic about the potential for older adults to continue their lifetime of learning, as is increasingly required in our rapidly-changing
technological environment. It is often assumed that crystalized intelligence is “merely” the preservation of established knowledge, and does not help to cope with novel situations. However, the present findings demonstrate that strong verbal knowledge enables older adults to comprehend unfamiliar and complex metaphors—their established knowledge can be applied to expressions they had never previously encountered. This finding aligns with work showing that older adults use semantic knowledge to aid in comprehension and memory in other domains (Artuso & Belacchi, 2021; Milburn et al., 2021; Umanath & Marsh, 2014). A practical implication of our findings is that metaphors (verbal and perhaps also non-verbal) may provide valuable tools for teaching older adults new concepts and skills, such as how to use novel technological devices (e.g., Micocci & Spinelli, 2018).

We note that the present study dealt only with the comprehension of metaphors, and not the use or creation of metaphorical thought. Much less is known about the more complex processes by which novel metaphors can be created by poets and others, and how this may change with age. At least in a few individuals, the capacity to generate metaphors survives into old age. A few days before he died at age 74, the great Irish poet Seamus Heaney composed “Banks of the Canal”, which opens with a slow-moving stream of metaphors.

Say ‘canal’ and there’s that final vowel
Towing silence with it, slowing time
To a walking pace, a path, a whitewashed gleam
Of dwellings at the skyline. World stands still.
References


https://doi.org/10.1037/bul0000145


DOI: 10.1080/13825585.2021.1962791


https://doi.org/10.1037/a0019578.


https://doi.org/10.1016/j.dr.2020.100925


https://psycnet.apa.org/doi/10.1080/03610730290080317


