

# Novel metaphor comprehension: Semantic neighbourhood density interacts with concreteness

Hamad Al-Azary<sup>1</sup> · Lori Buchanan<sup>2</sup>

Published online: 6 September 2016  
© Psychonomic Society, Inc. 2016

**Abstract** Previous research suggests that metaphor comprehension is affected both by the concreteness of the topic and vehicle and their semantic neighbours (Kintsch, 2000; Xu, 2010). However, studies have yet to manipulate these 2 variables simultaneously. To that end, we composed novel metaphors manipulated on topic concreteness and semantic neighbourhood density (SND) of topic and vehicle. In Experiment 1, participants rated the metaphors on the suitability (e.g. sensibility) of their topic-vehicle pairings. Topic concreteness interacted with SND such that participants rated metaphors from sparse semantic spaces to be more sensible than those from dense semantic spaces and preferred abstract topics over concrete topics only for metaphors from dense semantic spaces. In Experiments 2 and 3, we used presentation deadlines and found that topic concreteness and SND affect the online processing stages associated with metaphor comprehension. We discuss how the results are aligned with established psycholinguistic models of metaphor comprehension.

**Keywords** Language comprehension · Psycholinguistics · Semantics · Metaphor processing

In metaphor, unrelated concepts are paired to create a meaningful relation. For example, in *time is money*, the topic, or first concept, *time* and the vehicle, or second

concept, *money*, are semantically distinct but the statement is nonetheless comprehensible even though literally untrue. Given that the topic and vehicle are unrelated, they each refer to a wide range of properties which vary in their relevance to the created meaning of the metaphor. Understanding a metaphor requires emphasizing the properties relevant to its meaning while suppressing the irrelevant ones (Black, 1962). For example, when reading a metaphor such as *some lawyers are sharks*, irrelevant property primes (e.g. *sharks can be blue*) hinder processing whereas relevant property primes (e.g. *sharks can be ruthless*) facilitate processing, as compared to no primes (McGlone & Manfredi, 2001). Moreover, metaphor primes (e.g. *that defense lawyer is a shark*) facilitate reading target sentences that refer to properties which are relevant to the metaphor's meaning (e.g. *sharks are tenacious*) but not sentences that refer to properties that are irrelevant to the metaphor's meaning (e.g. *sharks are good swimmers*; Gernsbacher, Keysar, Robertson, & Werner, 2001). Taken together, the aforementioned studies suggest that online metaphor comprehension hinges on accessing the appropriate sense of the vehicle. Psycholinguistic models characterize the semantic processing of the topic and vehicle needed for metaphor comprehension (see Gibbs & Colston, 2012, for a review of such models). Of these models, one of the most comprehensive is Kintsch's (2000, 2001, 2008) predication algorithm, which we will discuss below.

The predication algorithm models the meaning of argument-predicate sentences, of which metaphor is an example, in semantic space (Kintsch, 2001). In this approach, latent semantic analysis (LSA) is used to model semantic space based on the co-occurrence of words in natural language (see Landauer, Foltz, & Laham, 1998, for an introduction to LSA). LSA assumes words that

✉ Hamad Al-Azary  
halazary@uwo.ca

<sup>1</sup> Department of Psychology, University of Western Ontario, London, ON, Canada

<sup>2</sup> University of Windsor, Windsor, ON, Canada

appear in similar contexts throughout written language are semantically related. Moreover, because in LSA words are represented in semantic space as vectors, similarity between them is inferred by their respective cosine. For example, according to LSA, the five most semantically similar words to *eye* are *cornea* (cosine: 0.82), *retina* (cosine: 0.82), *eyeball* (cosine: 0.81), *iris* (cosine: 0.79), and *sclera* (cosine: 0.73).<sup>1</sup> Words that are similar in meaning are described as *semantic neighbours* because they are closer in semantic space than words which are more dissimilar in meaning. To model the meaning of a sentence, the predication algorithm computes a vector based on the semantic neighbours of the predicate, along with the argument of the sentence. Although the predication algorithm is a general model originally developed for literal language, it can be applied to modelling metaphor comprehension as we will outline below (see Kintsch, 2001, for a demonstration of the predication algorithm in modelling other semantic processes).

To model the meaning of a metaphor, the predication algorithm constructs a spreading activation network that searches the semantic neighbourhood of the vehicle (i.e. predicate) for words that also happen to be related to the topic (i.e. argument), while inhibiting those that are unrelated. That is, the neighbours that are related to both the topic and the vehicle are strengthened whereas the other neighbours of the individual words are inhibited. The semantic neighbours that are activated (i.e. the semantic neighbours of the vehicle which are also related to the topic) along with the topic and vehicle are then used to compute a vector which is taken to represent the metaphor's meaning. This vector is the centroid of the topic, vehicle, and those semantic neighbours of the vehicle which are also related to the topic. The metaphor vector can be checked against words which should be related to its meaning. For example, the vector of the metaphor, *my lawyer is a shark* is highly related to the vector of the word *vicious* and much less related to the vector of the word *fish*,<sup>2</sup> which is what one would expect given the metaphor's meaning (Kintsch, 2000, 2001). Therefore, the predication algorithm appropriately selects the properties which are related to a metaphor's meaning, and inhibits those which are unrelated. Simulations of the predication algorithm have also been successfully tested against human interpretations of novel metaphors (Kintsch & Bowles, 2002).

Although robust, the predication algorithm leaves key questions about metaphor comprehension unanswered. First, it does not test the influences (if any) of word concreteness. Concreteness is an intriguing variable

considering that many metaphors employ an abstract topic and concrete vehicle (e.g. Gentner, 1983; Katz, 1989; Kintsch, 2000; Lakoff & Johnson, 1980; Wolff & Gentner, 2011; Xu, 2010). However, it is unclear if metaphors are indeed more comprehensible if they have an abstract rather than concrete topic. On the one hand, Xu (2010) reported that participants judged abstract topics and concrete vehicles (e.g. *revolution is an earthquake*) to be more similar to one another than concrete topics and concrete vehicles (e.g. *a pen is a sword*). On the other hand, Harris, Friel, and Mickelson (2006) found the opposite effect; in their study, participants rated metaphors made up of concrete topics and vehicles (e.g. *the hungry mosquitos were vampires*) to be more suitable for the discourse goal of comparing similarities than metaphors made up of abstract topics and concrete vehicles (e.g. *the harsh criticism was a stinging bullet*). Therefore, the effects of topic concreteness (or abstractness) on metaphor comprehension are still unclear.

Second, although the predication algorithm demonstrates that semantic neighbourhoods are implicated in metaphor comprehension, it does not describe those neighbourhoods in a way that predicts the degree to which a metaphor is comprehensible. Kintsch (2000) speculated that many semantic neighbours would be conducive to processing: "What strong metaphors seem to have in common is that the predicate is a concrete term, rich in imagery and many potential associations" (p. 261). However, he never tested if the density of neighbours affects a metaphor's comprehensibility or aptness. Katz (1992) also suggested that concrete vehicles are superior to abstract vehicles because the former are from denser semantic spaces (measured by word association tasks) and provide more semantic information. These speculations regarding the role semantic density plays in metaphor comprehension have yet to be tested.

## The present study

Kintsch's research suggests that semantic neighbourhoods of the topic and vehicle are involved in processing novel metaphors, and the aforementioned research points to topic concreteness as a crucial consideration. Interestingly, Paivio and Walsh (1993) note that the verbal associates of the topic and vehicle, along with their capacity to elicit imagery via concreteness, must both be considered for a full model of metaphor comprehension. As far as we know, there are no data upon which such a model can be based. To that end, in the following experiments we manipulate semantic neighbourhoods and topic concreteness in novel metaphor-comprehension tasks to examine their conjoint effects.

<sup>1</sup> These cosine values are derived from the LSA home-page (<http://lsa.colorado.edu/>).

<sup>2</sup> These observations are produced by the predication algorithm which, as far as we know, is not available on the LSA home page.

## Experiment 1

The goal of Experiment 1 was to determine if topic concreteness and general semantic neighbourhood characteristics both contribute to novel metaphor comprehension in an offline rating task. We manipulated the semantic neighbourhood density (SND) of the topics and vehicles in metaphors. SND was derived from the WINDSORS global word co-occurrence database and is a measure of how many near semantic neighbours a word has (Durda & Buchanan, 2008).

The WINDSORS measure of SND has been recently tested in some psycholinguistic tasks. For instance, Danguécan and Buchanan (2016) found an inhibitory effect from near neighbours in a lexical decision task; words from dense semantic spaces, or high SND words, were processed slower than words from sparse semantic spaces, or low SND words. MacDonald (2013) replicated the inhibitory effect in both young (18–25 years old) and older (60–80 years old) adults. Finally, McHugh and Buchanan (2016) found that WINDSORS semantic distances reflect the dominant and subordinate meanings of homographic words. In the WINDSORS database, a target word such as *depression* is more closely related to its dominant meaning, such as *sadness* than its subordinate meaning such as *hole*. Priming the dominant meaning (e.g. *sadness*) of a target word (e.g. *depression*) resulted in faster recognition than priming the subordinate meaning (e.g. *hole*) of the same target word. In summary, the previous studies that used semantic characteristics derived from the WINDSORS model all found that it characterizes semantic density in a way that is consistent with our current understanding of semantic processing.

Although no study has investigated the effects of SND on metaphor, given Kintsch's (2000) and Katz's (1992) speculations, we expect that metaphors made up of words from dense semantic spaces (i.e. high SND) will be rated more comprehensible than those made up of words from sparse semantic spaces (i.e. low SND). We also expect, given the concreteness literature reviewed above, that abstract-topic metaphors will be more comprehensible than concrete-topic metaphors. Finally, given that this is the first metaphor comprehension study to manipulate concreteness and SND, we have no basis for a prediction of an interaction.

## Method

### Participants

Fifty-two participants from the University of Windsor Psychology Participant Pool participated for extra course credit. All were 18 years of age or older, with normal or

corrected-to normal-vision, and were fluent English speakers.

### Materials

Nominal novel metaphors which juxtapose two nouns (e.g. *language is a bridge*) were developed to meet four experimental conditions which varied on topic concreteness and SND. Concreteness was defined as a dichotomous variable so that a concrete word refers to a tangible concept (e.g. *pen*) whereas an abstract word does not (e.g. *language*). Topic concreteness was varied so that metaphors either had an abstract or concrete topic, whereas vehicles remained concrete. Furthermore, whole metaphors were manipulated on WINDSORS SND so that both topic and vehicle were either high SND or low SND. For example, a high SND metaphor contains both a high SND topic (e.g. *pen*) and vehicle (e.g. *sword*) while a low SND metaphor contains both a low SND topic (e.g. *library*) and vehicle (e.g. *sanctuary*). High and low SND cut-offs were taken from Danguécan and Buchanan's (2016) visual word recognition study. Thus the experimental conditions were abstract topic–concrete vehicle, high SND (abstract-high SND); abstract-topic–concrete vehicle, low SND (abstract-low SND); concrete topic–concrete vehicle, high SND (concrete-high SND); and concrete topic–concrete vehicle, low SND (concrete-low SND). Some of the metaphors were borrowed or modified by others' work (Danguécan & Buchanan, 2016; Katz, Paivio, Marschark, & Clark, 1988; Xu, 2010). However, it was difficult to find or create many metaphors which would fit with our conditions; consequently, only 12 items were used per condition.

Nonsense fillers were used to compare their comprehensibility relative to the metaphors. Such statements were matched to the metaphor condition in concreteness and SND. The only difference is that during their construction they were intended to be meaningless (i.e. *imagination is a square*). Literal fillers (i.e. *a gorilla is an ape*) were also included and their SND levels were manipulated, but concreteness was not, because creating abstract literal statements proved to be a difficult task. Three practice items were also used which were not subject to any analysis. In total, 120 statements were used (not including three practice statements); 48 metaphors, 24 literal statements, and 48 nonsense statements. All items, along with on-screen task instructions, are shown in Appendix 1.

### Procedure

The entire task was undertaken on a computer with DirectRT (Jarvis, 2006) software used to present the

stimuli and collect responses. Participants were instructed both on-screen and orally to rate from 1 to 6 how suitable/natural statements seemed. Each participant rated statements from all conditions presented in a random order. The experiment was typically completed in less than 15 minutes.

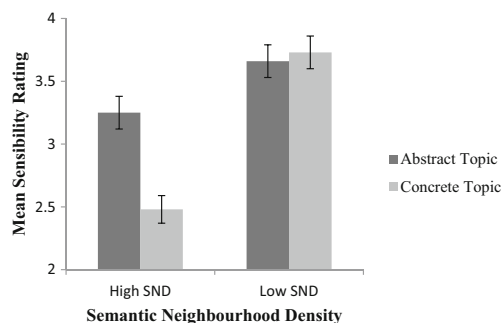
## Results

Data were analyzed by repeated measures analysis of variance (ANOVA). First, a one-way ANOVA was used to compare statement type (literal vs. metaphor vs. nonsense). As expected, we found a main effect of statement type,  $F(2, 102) = 1,166, p < .001, \eta^2 = .96$ . Follow-up  $t$ -tests revealed that metaphors ( $M = 3.28, SE = .11$ ) were more comprehensible than nonsense statements ( $M = 1.44, SE = 0.05$ ),  $t(51) = -23.11, p < .001$ , and less comprehensible than literal statements ( $M = 5.7, SE = .04$ ),  $t(51) = 20.53, p < .001$ .

The metaphors were analyzed by condition with a 2 (abstract topic vs. concrete topic) by 2 (high SND vs. low SND) repeated-measures ANOVA. A main effect of concreteness was obtained,  $F(1, 51) = 19.32, p < .001, \eta^2 = .28$ . Ignoring SND, abstract-topic metaphors ( $M = 3.45, SE = .12$ ) were rated as more sensible than concrete-topic metaphors ( $M = 3.11, SE = .11$ ). Moreover, a main effect of SND was obtained,  $F(1, 51) = 166.39, p < .001, \eta^2 = .77$ , but this effect was not in the predicted direction: Low SND metaphors ( $M = 3.70, SE = .12$ ) were rated as more sensible than high SND metaphors ( $M = 2.86, SE = .11$ ). More important, a significant concreteness by SND interaction,  $F(1, 51) = 41.96, p < .001, \eta^2 = .45$ , revealed that the difference in topic-concreteness sensibility ratings was present for high SND metaphors but not for low SND metaphors. See Fig. 1 for this interaction.

## Discussion

The main effect of SND is in contrast with the previous speculations of Katz (1992) and Kintsch (2000) that



**Fig. 1** Concreteness by SND interaction in Experiment 1. Error bars represent standard error of the mean

semantic density is conducive to metaphor comprehension. Recall that the predication algorithm searches the vehicle's semantic neighbourhood for items which are related to the topic. It seems that this search may be disrupted when the topic and vehicle have many near neighbours. On the other hand, when the topic and vehicle have fewer near neighbours, finding the appropriate ones can be carried out more effectively, resulting in a more comprehensible metaphor. Similarly, if metaphor comprehension is achieved in part by suppressing irrelevant properties, as Black (1962) reasoned, dense semantic neighbourhoods may contain more irrelevant properties and hence require more suppression than sparse semantic neighbourhoods.

The obtained interaction suggests that abstract-topic metaphors are more comprehensible than concrete-topic metaphors only when they are from dense semantic spaces, or high SND. Conversely, when the topic and vehicle are from sparse semantic spaces, or low SND, there appears to be no effect of topic concreteness. These results demonstrate the need to control for variables like SND when conducting metaphor-comprehension research. As previously discussed, Xu (2010) reported that participants rated abstract topic and concrete vehicles to be more similar to one another than concrete topics and concrete vehicles. Conversely, Harris et al. (2006) found the opposite effect; concrete topic metaphors were judged to suggest similarity more than abstract topic metaphors. The discrepancy in those studies' results could be because SND was not controlled for.

Although there were differences in sensibility ratings, it is not clear if the aforementioned task measured actual *comprehension*; it could be that the open-ended sensibility task actually measured metaphor aptness, which does not necessarily reflect metaphor comprehension (Gerrig & Healy, 1983). That is, after comprehending the metaphors in a similar way, participants may have considered their aptness before making sensibility judgements. Moreover, comprehensibility ratings (presumably similar to the sensibility ratings used here) and aptness are positively correlated (as reviewed by Jones & Estes, 2006). To get a clearer picture of comprehension processes, in Experiment 2 we used the same stimuli in a task which is thought to characterize the online comprehension stages of metaphors.

Wolff and Gentner (2011) recently pinpointed the processing stages associated with the time course of metaphor comprehension. In their clever task, participants read metaphors (e.g. *some suburbs are parasites*) and reversed metaphors (e.g. *some parasites are suburbs*) at 600- and 1,600-ms deadlines. After reading any given metaphor, participants rapidly (within 400 ms) rated the statements as comprehensible or incomprehensible. At the 600-ms deadline, metaphors in the reversed form were

rated just as comprehensible as their forward counterparts (mean comprehension scores for forward and reversed metaphors were .39 and .36, respectively). Conversely, at the later deadline of 1,600 ms, forward metaphors increased in comprehension ratings whereas reversed metaphors did not (mean comprehension scores for forward and reversed metaphors were .54 and .35, respectively). Nonsense controls were rated lower than metaphors (including reversed metaphors) on comprehensibility at both presentation deadlines whereas literal controls were rated higher than metaphors. Wolff and Gentner (2011) concluded that the early stage, in which participants did not show a preference for the orientation of the metaphor, indicates *symmetrical* processing, in that shared features between the topic and vehicle are bidirectionally mapped, and the later processing stage is *asymmetrical*, in that relations in the vehicle are projected to the topic (also see Gentner & Bowdle, 2008). Therefore, Wolff and Gentner (2011) showed that (at least) two qualitatively different processing stages, the first occurring by 600 ms and the second by 1,600 ms, underlie metaphor comprehension.

In Experiment 2, we used Wolff and Gentner's (2011) presentation deadlines of 600 (early stage) and 1,600 (late stage) ms to further study the effects of SND and concreteness on metaphor comprehension. We expect that at 600 ms, comprehension ratings for metaphors will be equal. Our rationale for this prediction is that Wolff and Gentner (2011) revealed that this processing stage is too shallow for participants to be sensitive to the reversal of topics and vehicles; therefore, we anticipate it will also be too shallow for word-level semantic effects. Regarding the late processing stage, we predict that if the concrete-high SND metaphors are indeed less comprehensible than the other metaphors, as suggested by the results of Experiment 1, they should not increase in comprehension ratings by the late stage. The low SND, and the abstract-high SND metaphors, however, should increase in comprehension ratings by the late stage, which would indicate that they are more comprehensible and "reach" the later stage of processing.

## Experiment 2

### Method

#### Participants

Fifty people participated for partial course credit. Recruitment and sample characteristics were the same as in Experiment 1.

### Materials and procedure

The same stimuli from Experiment 1 were used, with the addition of more practice items (see below). Participants were provided with instructions on-screen (see Appendix 2 for instructions and stimuli) and orally informed that they would be quickly judging the comprehensibility of statements using a button box. Participants were encouraged to dedicate their left hand for the button on the left side, to be pressed if a statement were incomprehensible, and their right hand for the button on the right side, to be pressed if a statement were comprehensible. Similar to Wolff and Gentner's (2011) procedure, a practice session was initiated to orient the participants to the buttons and their corresponding representations. This practice session consisted of presenting the words *comprehensible* and *incomprehensible* at the same presentation deadlines as the experimental items to come. Words were preceded by a 300-ms presentation of pound signals which matched the number of letters in each word. The words were presented for 600 ms in one condition and in another for 1,600 ms. A response-terminated question mark followed each presentation. In short, the stimulus presentation schedule was pound signals for 300 ms, replaced by the word for either 600 or 1,600 ms, replaced by a response-terminated question mark. Participants were instructed to make a response at the sight of the question mark and were told that they only had a limited amount of time to respond. An error message reading *Please try to respond faster!* appeared after any trial in which a response was made 400 ms after the presentation of the question mark. In total, this practice session had 20 trials. Half of the trials had the word *comprehensible* presented at both presentation durations whereas the other half had the word *incomprehensible* presented at both presentation durations. The correct responses for the *comprehensible* and *incomprehensible* words were the right and left buttons, respectively, pushed within 400 ms of the presentation of the question mark. Participants were supervised and feedback was provided during this session. At the conclusion, another practice session was initiated. This practice session was identical in its stimulus presentation schedule; however, statements (15 nonsense, 15 metaphor, and 15 literal) were used in place of the single words. Participants were reminded to press the button on the right if the statement was comprehensible and the button on the left if the statement was incomprehensible. In total, this practice session involved 90 trials. After this session, the experimenter left the testing room, and the testing session was initiated with 240 experimental trials presented in a random order (Experiment 1 stimuli presented at both processing deadlines). Participants finished the entire study in less than 30 minutes.

## Results

Following Wolff and Gentner's (2011) data clean-up procedure, we removed from the analyses any responses made after 400 ms of the question mark presentation. This was to ensure that all participants remained on task and did not process the metaphors beyond their allotted deadlines. This resulted in the removal of 17.3 % of the trials, which is less than the percentage of trials removed in Wolff and Gentner's (2011) experiment with the same presentation deadlines. This difference may be attributed to the use of the question mark, which was a signal for participants to make their comprehensibility judgements. Wolff and Gentner (2011) did not use a question mark, and as a result their participants may have been more likely to miss inputting their comprehensibility judgements on time. See Appendix 3 for a breakdown of data cleanup for each condition.

### Statement comprehension

Before the main analysis, the effect of statement type was examined at both deadlines to ensure that participants were interpreting metaphors as more comprehensible than nonsense statements and less comprehensible than literal statements. This was achieved by a statement (nonsense vs. metaphor vs. literal) by deadline (early vs. late) repeated-measures ANOVA. A main effect of statement was obtained,  $F(2, 98) = 498.12, p < .001, \eta^2 = .91$ , as well as a main effect of deadline,  $F(1, 49) = 15.50, p < .001, \eta^2 = .24$ . Moreover, a statement by deadline interaction was obtained,  $F(2, 98) = 76.48, p < .001, \eta^2 = .61$ . Bonferroni adjusted  $t$ -tests revealed a significant difference between metaphors and literals at the early deadline,  $t(49) = 13.21, p < .001$ , and the later deadline,  $t(49) = 17.78, p < .001$ . This was also true for metaphors and nonsense statements at the early deadline,  $t(49) = 10.27, p < .001$ , and the late deadline,  $t(49) = 17.22, p < .001$ .

These results confirm that participants recognized novel metaphors as more meaningful than nonsense statements but less meaningful than literal statements. Furthermore, participants were not simply guessing when making their comprehension judgements because guessing should result in even comprehension ratings for each statement type at each processing deadline.

### Main analysis of metaphors

A concreteness by SND by deadline repeated-measures ANOVA revealed a main effect of concreteness,  $F(1, 49) = 7.83, p = .007, \eta^2 = .14$ . Overall, metaphors made up of abstract topics ( $M = .48, SE = .04$ ) were more comprehensible than those made up of concrete topics ( $M = .43, SE = .03$ ). A main effect of SND was obtained,  $F(1, 49) = 52.78, p <$

$.001, \eta^2 = .52$ . Metaphors made up of low-SND words ( $M = .50, SE = .02$ ) were more comprehensible than their high-SND counterparts ( $M = .45, SE = .02$ ). Finally, a main effect of deadline was obtained,  $F(1, 49) = 16.06, p = .001, \eta^2 = .25$ . Overall, metaphors presented at the later processing stage ( $M = .49, SE = .01$ ) were more comprehensible than metaphors presented at the early processing stage ( $M = .46, SE = .02$ ).

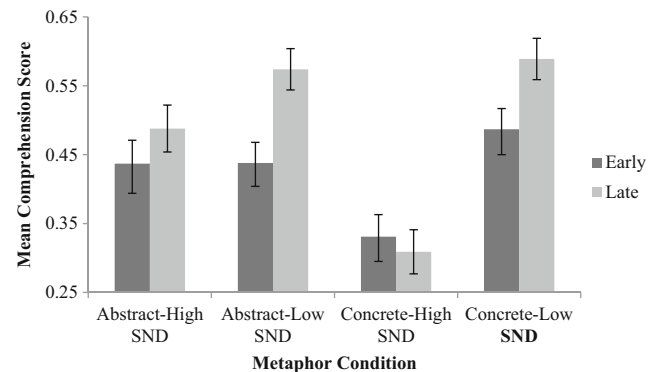
Several interaction effects were revealed, including a concreteness by SND interaction,  $F(1, 49) = 39.00, p < .001, \eta^2 = .44$ ; like in Experiment 1, there was a difference between abstract-high SND ( $M = .46, SE = .04$ ) and concrete-high SND metaphors ( $M = .32, SE = .03$ ) but not between abstract-low SND ( $M = .51, SE = .03$ ) and concrete-low SND ( $M = .54, SE = .03$ ) metaphors. Furthermore, a concreteness by deadline interaction was obtained,  $F(1, 49) = 5.17, p = .027, \eta^2 = .10$ ; there was a comprehension difference between abstract metaphors at the early ( $M = .44, SE = .04$ ) and late ( $M = .53, SE = .04$ ) deadlines, but not between concrete metaphors at the early ( $M = .41, SE = .03$ ) and late ( $M = .45, SE = .03$ ) deadlines.

Finally, a semantic neighborhood density by deadline interaction was obtained,  $F(1, 49) = 17.55, p < .001, \eta^2 = .26$ . The effects of SND on comprehension varied across levels of deadline. There was a difference in comprehensibility ratings between low-SND metaphors at the early ( $M = .46, SE = .03$ ) and late ( $M = .58, SE = .03$ ) presentation deadlines, but not between high SND metaphors at the early ( $M = .38, SE = .04$ ) and late ( $M = .40, SE = .03$ ) deadlines.

A three-way interaction was nonsignificant,  $F(1, 49) = 1.17, p = .46, \eta^2 = .01$ . Figure 2 shows each of the condition means at both processing deadlines. Concrete-high SND metaphors, as predicted, do not increase in comprehension at the later stage of processing.

## Discussion

One striking difference between this set of results and Wolff and Gentner's (2011) is that at the early processing stage,



**Fig. 2** Mean comprehension score for each of the metaphoric conditions at both processing deadlines for Experiment 2. Error bars represent standard error of the mean

concrete-high SND metaphors were distinguishable from the other conditions, as shown by their lower ratings. Recall that Wolff and Gentner (2011) found the early processing stage of 600 ms to be too short for participants to distinguish between forward or reversed metaphors. The fact that the current stimulus set yields a comprehension difference among forward metaphor types at 600 ms of processing time is surprising and is a testament to the robust effects of SND and concreteness.

The obtained results further demonstrate that there are comprehensibility differences between metaphors manipulated on topic concreteness and SND. Unlike in Experiment 1, the results obtained here demonstrate that there are differences between the online processing of the metaphors. There is, however, a potential confound in the experimental design that needs to be addressed. Recall that we employed a repeated-measures design, so participants were exposed to each metaphor at both deadlines. Although the reoccurrence of each metaphor was random, there is, nonetheless, a potential response bias whereby participants base a proportion of their responses to metaphors presented the second time in the list on their exposure to them the first time on the list. This possibility is especially relevant because Bowdle and Gentner (2005) found comprehension differences between metaphors briefly taught to participants and metaphors that were never before encountered. Therefore, to be sure that there is no exposure effect, we will eliminate this potential bias by replicating Experiment 2 with presentation deadline as a between-participants variable.

### Experiment 3

#### Method

##### *Participants, materials, and procedure*

Seventy-one people participated for partial course credit. Recruitment and sample characteristics were the same as in Experiments 1 and 2. The experimental procedures and stimuli were identical to Experiment 2. The only procedural difference was, unlike in Experiment 2, where participants saw the same metaphors at both the early and late presentations, here deadline was implemented as a between-participant variable; 37 participants viewed stimuli for 600 ms, whereas 34 participants viewed stimuli for 1,600 ms.

##### *Results*

Data removal followed the same procedures as outlined in Experiment 2; this resulted in the removal of 17.8 % of the data. (See Appendix 3 for a breakdown of trials removed by each condition.) One participant from the 1,600-ms condition

was removed from data analysis because they failed to respond within 400 ms in all of the statements of a given condition. Therefore, data were analyzed from 70 participants; 37 participants viewed stimuli for 600 ms and 33 participants viewed stimuli for 1,600 ms.

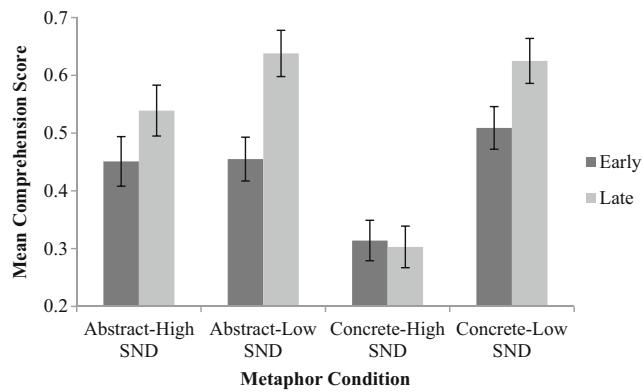
##### *Statement comprehension*

As in Experiment 2, a statement (nonsense vs. metaphoric vs literal) by deadline (early vs. late) mixed design ANOVA was run (with Greenhouse–Geisser correction). A main effect of statement was again obtained,  $F(2, 122.77) = 699.35, p < .001, \eta^2 = .91$ . Comparisons revealed that, similar to Experiment 2, concrete metaphors ( $M = .48, SE = .02$ ) were less comprehensible than literal concrete statements ( $M = .85, SE = .01$ ), but more comprehensible than nonsense statements ( $M = .19, SE = .01, p < .001$ ). A main effect of deadline approached significance;  $F(1, 68) = 3.96, p = .051, \eta^2 = .06$ . A statement by deadline interaction was obtained,  $F(1.8, 122.77) = 68.5, p < .001, \eta^2 = .50$ , and is in the same direction as Experiment 2.

##### *Main analysis and discussion*

As in the previous Experiments, only the metaphoric statements were subject to further analysis. A concreteness by SND by deadline mixed-design ANOVA revealed a main effect of concreteness  $F(1, 68) = 17.32, p < .001, \eta^2 = .20$ . Metaphors containing abstract topics ( $M = .52, SE = .03$ ) were rated as more comprehensible than metaphors containing concrete topics ( $M = .44, SE = .02$ ), which is consistent with Experiments 1 and 2. Also, a main effect of SND was obtained,  $F(1, 68) = 76.24, p < .001, \eta^2 = .53$ , and this too is consistent with Experiments 1 and 2. Metaphors made up of low-SND words ( $M = .56, SE = .02$ ) were rated as more comprehensible than metaphors made up of high-SND words ( $M = .40, SE = .02$ ). A between-subjects effect of deadline was obtained,  $F(1, 68) = 4.20, p = .04, \eta^2 = .06$ . Metaphors presented at the early presentation deadline of 600 ms ( $M = .43, SE = .03$ ) were less comprehensible than those presented at the later deadline of 1,600 ms ( $M = .53, SE = .03$ ).

The same interaction effects as those found in Experiment 2 were obtained again. A concreteness by deadline interaction was significant,  $F(1, 68) = 4.81, p = .032, \eta^2 = .07$ , as was the SND by deadline interaction,  $F(1, 68) = 9.86, p = .003, \eta^2 = .13$ . Furthermore, a concreteness by SND interaction was also obtained,  $F(1, 68) = 31.54, p < .001, \eta^2 = .32$ . These interactions were both in the same directions as Experiment 1. The three-way interaction, like in Experiment 2, was nonsignificant,  $F(1, 68) = .172, p = .680, \eta^2 = 0.003$ . Figure 3 shows each of the condition means at both processing deadlines. As can be seen, metaphors made up of low-SND words appear to benefit, as shown by increased comprehension ratings, from later



**Fig. 3** Mean comprehension score for each of the metaphoric conditions at both processing deadlines for Experiment 3. Error bars represent standard error of the mean

processing deadlines. Metaphors made up of high-SND words and with concrete topics do not result in increased comprehension ratings as a result of later processing deadlines.

To summarize, this replication of Experiment 2 resulted in the same pattern of findings and rules out the possibility that response bias or stimulus familiarity could have produced the effects of interest.

## General discussion and future directions

The experiments reported here are the first to simultaneously manipulate topic concreteness along with SND in novel metaphor comprehension. The main finding from Experiment 1 is that SND interacted with topic concreteness such that low-SND metaphors, regardless of topic concreteness, are equally comprehensible, but concrete-high SND metaphors are less comprehensible than abstract-high SND metaphors. Moreover, the SND effect in general is inconsistent with Kintsch's (2000) and Katz's (1992) speculation on semantic density being conducive to comprehension. Regarding the predication algorithm, it is unclear why many near neighbours would have a detrimental effect on computing a metaphor's meaning, except if we extend the model to suggest that density increases the complexity of determining shared neighbours between topic and vehicle. Perhaps an apt metaphor to describe our findings regarding semantic density is that less is more.

Experiments 2 and 3 borrowed from Wolff and Gentner's (2011) methodology using presentation deadlines. In fact, the goal of their study was to demonstrate the processing stages of the structure-mapping model (see Gentner, 1983; Gentner & Bowdle, 2008; Wolff & Gentner, 2011). Structure-mapping holds that metaphor comprehension is achieved by two stages which compare the topic and vehicle. In the first stage, structures shared by the topic and vehicle are symmetrically aligned, whereas in the second stage, structures in the vehicle are

asymmetrically projected to the topic. For example, the metaphor *some suburbs are parasites* is understood by aligning the shared relation between suburbs and parasites, namely, *existence in dependence of a host* and then by projecting a vehicle-specific relation to the topic, such as *harms its host* (Wolff & Gentner, 2011). As we have reviewed, Wolff and Gentner (2011) found support for structure-mapping by showing that metaphors are processed symmetrically at 600 ms and asymmetrically at 1,600 ms. Structure-mapping would predict that all metaphors increase in comprehension by the late stage of 1,600 ms. Accordingly, this is when vehicle-specific relations are projected to the topic. The fact that the metaphors from the three superior conditions in our experiments (i.e. abstract-high SND, abstract-low SND, and concrete-low SND) increased in comprehension by the late presentation deadline is consistent with Wolff and Gentner's (2011) findings. However, lack of change in comprehension in the concrete-high SND metaphors is inconsistent with their claims. It is unclear why the semantic density of concrete-high SND metaphors can disrupt the structure-mapping process.

The results may fit with other accounts of metaphor processing, such as the quality-of-metaphor hypothesis (Glucksberg, 2008; Glucksberg & Haught, 2006). This view argues that metaphor aptness determines how a metaphor is processed. Apt metaphors are processed by categorizing the topic into a superordinate category in which the vehicle stands in for. For example, *my lawyer is a shark* can be understood by constructing a metaphorical category that shark belongs to, such as *vicious things* (Glucksberg, 2008). However, if a metaphor is poor, it is processed as a comparison in which shared features of the topic and vehicle are matched, similar to structure-mapping (Glucksberg & Haught, 2006). It seems that our results can supplement the quality-of-metaphor hypothesis by defining what makes a metaphor good or poor.

In fact, a shortcoming of the quality-of-metaphor hypothesis is that it does not define what a good or poor quality metaphor looks like. As Gentner and Bowdle (2008) pointed out, there is a flaw in models which suggest different processing mechanisms for figurative and literal language; namely, that a statement must be identified as figurative or literal for it to be processed accordingly, but identifying it entails processing it. We suggest a similar criticism is warranted for the circular reasoning of the quality-of-metaphor hypothesis. That is, a metaphor must be processed to determine its quality. Unless the quality-of-metaphor hypothesis specifies what makes a metaphor apt, it is unclear how any given metaphor is processed before its quality is known. However, word-level semantic characteristics, such as topic concreteness and SND, stored in the lexicon may identify metaphor



quality. The more comprehensible metaphors used in our study, as shown by their comprehensibility ratings, may be thought of as good metaphors, whereas the less comprehensible (i.e. concrete-high SND metaphors) may be considered poor metaphors. Therefore, semantic characteristics may define metaphor quality and, potentially, whether a metaphor is processed as a categorization or a comparison.

Our view follows from the quality-of-metaphor hypothesis (Glucksberg, 2008; Glucksberg & Haught, 2006) in that we believe that metaphor involves placing the topic in the metaphoric category of the vehicle. However, like Kintsch's (2000, 2008), we believe that the category can be operationalized as the semantic neighbourhood of the vehicle. We also take into consideration Black's (1962) insight that metaphor comprehension is necessitated by suppressing any properties of the topic and vehicle which are unrelated and hence may interfere with the meaning of the metaphor. We offer this more nuanced and qualitatively testable explanation: If a topic is placed in a semantic neighbourhood, then a dense neighbourhood may have too many associations and not enough room to assimilate a new word. On the other hand, sparse semantic spaces would have the room required to assimilate a new word. In the case of high-SND metaphors, topic concreteness matters because abstract words typically have less physical entities (Wiemer-Hastings & Xu, 2005). Thus, abstract words would fit in dense spaces better than concrete words. If concrete words have more attributes than abstract words, then categorizing concrete words would be more difficult in a dense neighbourhood because there are many close neighbours that must cohere with the concrete word and its features. For example, consider two high SND metaphors: *a pen is a sword* and *language is a bridge*. As we revealed, the latter is more comprehensible than the former. We believe this is because the lack of concrete features in *language* allows ease of assimilation into a dense neighbourhood. On the other hand, *pen* has many concrete features that impede categorizing it in the semantic neighbourhood of *sword*. Such metaphors may be processed as comparisons. In sparse spaces, however, concreteness is not such an issue; abstract and concrete topics should have equal or near equal assimilation.

As alluded to previously, the data presented here provide a potential explanation with respect to the concreteness effect in metaphor comprehension (e.g. Harris, Friel, & Mickelson, 2006; Xu, 2010). It may be that these inconsistencies arise because SND for the items for those studies was not considered and we have shown here that SND and concreteness interact. This consideration may add an interesting twist to a recent study that looked at concreteness in metaphors. Forgács, Bardolph, Amsel, DeLong, and Kutas (2015) studied how adjective-noun metaphors (e.g. *thin schedule*) compared to concrete literal (e.g. *printed schedule*) and abstract literal (e.g. *conditional schedule*)

statements in an event-related potential paradigm. The more concrete metaphors produced N400s that were similar to abstract literal statements whereas the more abstract metaphors produced N400s that were similar to concrete literal statements. However, considering that concreteness and SND interact, it could be that metaphors do not elicit different concreteness effects but that the stimuli used in their study were not matched on SND. Therefore, controlling for SND may clarify the inconsistent concreteness effects in the literature regarding topic-concreteness (e.g. Harris, Friel, & Mickelson, 2006; Xu, 2010) and N400s (e.g. Forgács et al., 2015).

Metaphor comprehension is affected by both bottom-up (e.g. word-level semantics) and top-down (e.g. context) processes (Burgess & Chiarello, 1996). Semantic neighbourhood density and concreteness characterize the bottom-up requirement; however, future research must consider how these semantic variables interact with a linguistic context. Moreover, one cannot overlook the idiosyncratic differences between people, such as education, occupation, and age, which may affect metaphor comprehension and interpretation (Gibbs, 2013). Although, we manipulated semantic variables of both topics and vehicles in novel metaphors, we are aware that the small number of items used in each condition, along with a narrow participant sample demographic, raises questions about the generalizability of our results, especially without considering context. Nonetheless, we consistently demonstrated that SND, paired with topic concreteness, is a valuable construct with broad but important theoretical implications and should be considered in future research. Such word-level semantic variables may help explain why some metaphors are better than others (or why some metaphors are difficult to understand; see Kintsch & Bowles, 2002), and what processing mechanisms are engaged during comprehension. Other important questions pertain to the comprehension difference between metaphors and similes (e.g. Haught, 2013, 2014) and why some metaphors are reversible and others are not (see Campbell & Katz, 2006). By continuing to consider word-level semantics, perhaps these questions can be addressed, and a fuller model of novel metaphor comprehension can be realized.

**Acknowledgments** The authors thank Albert Katz for comments on an earlier version of this manuscript.

## Appendix 1

### On-screen instructions for Experiment 1

Your task will be to rate how suitable or natural a series of statements are. The statements will either be nonsensical, literal, or figurative. For example, *A Sheep is a Hill* is a nonsensical statement; *A Circle is a Shape* is a literal statement; *Love is a Journey* is a figurative statement. Use the number pad to indicate your rating from 1 to 6

(1 being *very unsuitable/unnatural* and 6 being *very natural/suitable*). Please read the statements carefully before making a response. Press the space bar to begin. There will be three practice trials followed by the rest of the experiment. When you are finished, a thank-you message will appear. Afterwards, you may exit the room.

**Practice statements for Experiment 1:** A Circle is a Shape; A Sheep is a Hill; Love is a Journey

## Appendix 2

### On-screen instructions for Experiments 2 and 3

Your task will be to rate how comprehensible statements are. The statements will either be nonsensical, literal, or

#### Metaphors

Abstract High SND	Abstract Low SND
Civilization is a Crust	Censorship is a Filter
Language is a Bridge	Indecision is a Whirlpool
Superstition is a Disease	Austerity is a Remedy
Cynicism is an Undertaker	Daydream is a Trip
Addiction is Paste	Destiny is a Story
Justice is a Net	Digestion is a Bulldozer
Revolution is an Earthquake	Responsibility is a Chain
Heaven is Dessert	The Unconscious is a Factory
Passion is a Storm	Debate is a Pendulum
Ignorance is Blindness	Joy is Warmth
Revelation is Rain	Departure is a Sunset
Sarcasm is a Knife	Discovery is a Sunrise

Concrete High SND	Concrete Low SND
A Pen is a Sword	A Library is a Sanctuary
A Museum is a Cemetery	A Politician is a Broom
A Beach is a Grill	A Scarecrow is a Guardian
Embroidery is Ink	A Surfer is a Swan
A Tadpole is a Seed	Veins are Roots
A Mosquito is a Vampire	Lipstick is a Marker
A Cigarette is a Syringe	A Pond is a Mirror
A Cactus is a Bottle	A Woodpecker is a Lumberjack
A Crab is an Anchor	A Heart is a Motor
Money is Medicine	Darkness is a Cover
A Forest is a Harmonica	A Cloud is a Curtain
A Zebra is a Piano	A Star is a Sign

#### Nonsense Statements

Abstract High SND	Abstract Low SND
Depression is a Party	Imagination is a Square
Destruction is a Coat	Shelter is a Nose
Veneration is a Pickle	Patriotism is a Leaf
Philosophy is an Insect	Addition is a Beak
Religion is Snow	Tribute is a Stick
Argument is Paint	Depth is a Firework
Crime is a Raven	Confusion is an Alligator
Belief is a Reptile	Arrival is a Shoestring
Calculation is a Dinosaur	Exercise is a Roommate
Deception is a Cello	Sensation is a Suitcase
Evaluation is a Lamp	Suitability is a Donkey
Espionage is a Rock	Art is a Kitten

Concrete High SND	Concrete Low SND
A Table is a Fox	A University is a Spa
A Television is a Spear	A Bug is a Coin
A Theatre is a Bookshop	A Wallet is a Handkerchief
A Boot is a Brick	A Trunk is a Gear
A Cake is a Wrench	A Shell is a Sidewalk
A Kayak is a Spy	A Tooth is an Egg
A Bulldog is a Cherry	A Circus is a Pool
A Satellite is a Lightbulb	A Toe is a Coach
A Rabbit is a Pitcher	A Coast is a Tube
A Lizard is a Raindrop	A Napkin is Candy
A Fork is a Planet	A Staple is a Shelf
A Ladder is a Sailboat	An Armchair is a Script

#### Literal Statements

High SND

A Bus is a Vehicle  
 A Necklace is Jewellery  
 Banana is a Fruit  
 A Whale is a Mammal  
 A Frog is an Amphibian  
 Violet is a Colour  
 The Bible is Scripture  
 A Gorilla is an Ape  
 A Collie is a Pet  
 A Couch is Furniture  
 A Beard is Hair  
 A Poppy is a Flower

Low SND

Leather is a Material  
 A Screwdriver is a Tool  
 Chicken is a Meat  
 Juice is a Liquid  
 A Crocodile is a Predator  
 A Hamburger is Food  
 Cheddar is a Cheese  
 An Apartment is a Structure  
 A Turtle is a Creature  
 Gasoline is a Fuel  
 A Cannonball is a Sphere  
 A Mouse is a Rodent

Practice Statements for Experiments 2 and 3

Literal Statements	Metaphors	Nonsense Statements
A House is a Building	Adoration is Lightning	A Bar is a Wire
A Snake is Venomous	An Ambassador is a Peacock	A Tree is a Rocket
A Vegetable is Healthy	A Butterfly is a Flower	A Balloon is a Monkey
Facebook is a Website	A Dream is an Eclipse	A Stream is a Mountain
Golf is a Sport	Sleep is an Ocean	A Bug is an Orange
Chess is a Game	A Gene is a Blueprint	A Friend is a Scientist
A Skyscraper is Tall	Alcohol is a Crutch	A Battery is a Wrench
A Liver is an Organ	Lust is Anarchy	Grass is a Beaker
A Mammoth is Extinct	The Wind is an Arrow	Paper is a Trampoline
A Femur is a Bone	Cocaine is a Joyride	A Trailer is a Cup
Vision is a Sense	Truth is a Labyrinth	A Pear is an Animal
Spring is a Season	Happiness is Gold	A Sandwich is a Sauce
A Mushroom is a Fungus	Depression is a Ditch	A Feather is a Twig
Oak is a Wood	A Shadow is a Stalker	A Motorcycle is a Reptile
A Berry is a Fruit	A Baby is an Angel	A Park is a Jail

figurative. For example, *A Sheep is a Hill* is a nonsensical statement; *A Circle is a Shape* is a literal statement; *Love is a Journey* is a figurative statement. Treat metaphors, or figurative statements, as comprehensible. Use the button on the far right if the statement is comprehensible and the button on the far left if the statement is incomprehensible. Please wait until the statement disappears and a ? appears before making your response. Statements will be presented quickly, so please act as quickly and accurately as possible. The practice session will first involve the words comprehensible and incomprehensible. For comprehensible, press the button on the far right and for incomprehensible, press the button on the far left. Press the space bar to begin.

**On-screen instructions for practice session in Experiments 2 and 3**

Now you will do the same task but in response to statements. If the statement is comprehensible, press the button on the far right. If the statement is incomprehensible, press the button on the far left. Remember to react as quickly and accurately as possible. Press the spacebar to begin.

**On-screen instructions for experimental session in Experiments 2 and 3**

Now you will do the testing phase. Remember to respond as soon as you see the ?. Please press the space bar when you are ready.

**Appendix 3**

**Table 1** Percentage of responses removed from data analysis for Experiment 2

Condition	Metaphor	Nonsense Statement	Literal Statement
Abstract, High SND, Early	31%	28%	-
Abstract, High SND, Late	12%	7%	-
Abstract, Low SND, Early	27%	30%	-
Abstract, Low SND, Late	11%	8%	-
Concrete, High SND, Early	28%	24%	20%
Concrete, High SND, Late	8%	5%	6%
Concrete, Low SND, Early	31%	28%	22%
Concrete, Low SND, Late	9%	9%	6%

**Table 2** Percentage of responses removed from data analysis for Experiment 3

Condition	Metaphor	Nonsense Statement	Literal Statement
Abstract, High SND, Early	24%	23%	-
Abstract, High SND, Late	15%	12%	-
Abstract, Low SND, Early	23%	20%	-
Abstract, Low SND, Late	17%	16%	-
Concrete, High SND, Early	23%	18%	18%
Concrete, High SND, Late	15%	13%	10%
Concrete, Low SND, Early	22%	22%	18%
Concrete, Low SND, Late	17%	14%	10%

## References

- Black, M. (1962). *Models and metaphors* (Vol. 61, p. 1). Ithaca, NY: Cornell University Press.
- Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193–216.
- Burgess, C., & Chiarello, C. (1996). Neurocognitive mechanisms underlying metaphor comprehension and other figurative language. *Metaphor and Symbolic Activity*, *11*(1), 67–84.
- Campbell, J., & Katz, A. (2006). On reversing the topics and vehicles of metaphor. *Metaphor and Symbol*, *21*, 22–59.
- Danguécan, A. N., & Buchanan, L. (2016). Semantic neighborhood effects for abstract versus concrete words. *Frontiers in Psychology*, *7*. doi:10.3389/fpsyg.2016.01034
- Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods*, *40*, 705–712.
- Forgács, B., Bardolph, M. D., Amsel, B. B., DeLong, K. A., & Kutas, M. (2015). Metaphors are physical and abstract: ERPs to metaphorically modified nouns resemble ERPs to abstract language. *Frontiers in Human Neuroscience*. doi:10.3389/fnhum.2015.00028
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.
- Gentner, D., & Bowdle, B. (2008). Metaphor as structure-mapping. In R. W. Gibbs Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 109–128). Cambridge, UK: Cambridge University Press.
- Gernsbacher, M. A., Keysar, B., Robertson, R. R. W., & Werner, N. K. (2001). The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language*, *45*, 433–450.
- Gerrig, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 667.
- Gibbs, R. W., Jr. (2013). The real complexities of psycholinguistic research on metaphor. *Language Sciences*, *40*, 45–52.
- Gibbs, R. W., Jr., & Colston, H. L. (2012). *Interpreting figurative meaning*. Cambridge, UK: Cambridge University Press.
- Glucksberg, S. (2008). How metaphors create categories—Quickly. In R. W. Gibbs Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 67–83). Cambridge, UK: Cambridge University Press.
- Glucksberg, S., & Haught, C. (2006). On the relation between metaphor and simile: When comparison fails. *Mind & Language*, *21*, 360–378.
- Harris, R. J., Friel, B. M., & Mickelson, N. R. (2006). Attribution of discourse goals for using concrete- and abstract-tenor metaphors with or without discourse context. *Journal of Pragmatics*, *38*(6), 863–879.
- Haught, C. (2013). A tale of two tropes: How metaphor and simile differ. *Metaphor and Symbol*, *28*(4), 254–274.
- Haught, C. (2014). Spain is not Greece: How metaphors are understood. *Journal of Psycholinguistic Research*, *43*, 351–356.
- Jarvis, B. G. (2006). *DirectRT (Version 2006.2.0.28) [Computer software]*. New York, NY: Empirisoft Corporation.
- 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.
- Katz, A. N. (1989). On choosing the vehicles of metaphors: Referential concreteness, semantic distances, and individual differences. *Journal of Memory and Language*, *28*(4), 486–499.
- Katz, A. N. (1992). Psychological studies in metaphor processing: Extensions to the placement of terms in semantic space. *Poetics Today*, *13*(4), 607–632.
- 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*, 191–214.
- Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*(2), 257–266.
- Kintsch, W. (2001). Predication. *Cognitive Science*, *25*(2), 173–202.
- Kintsch, W. (2008). How the mind computes the meaning of metaphor. In R. W. Gibbs Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 129–142). Cambridge, UK: Cambridge University Press.
- Kintsch, W., & Bowles, A. R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, *17*(4), 249–262.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2/3), 259–284.
- Macdonald, G. (2013). *Aging and semantic processing*. Windsor, Ontario, Canada: Unpublished doctoral thesis, University of Windsor.
- McGlone, M. S., & Manfredi, D. A. (2001). Topic-vehicle interaction in metaphor comprehension. *Memory & Cognition*, *29*(8), 1209–1219.
- McHugh, T., & Buchanan, L. (2016). Pun processing from a psycholinguistic perspective: Introducing the model of psycholinguistic hemispheric incongruity laughter (M. PHIL). *Laterality: Asymmetries of Body, Brain and Cognition*, 1–29. doi:10.1080/1357650X.2016.1146292
- Paivio, A., & Walsh, M. (1993). Psychological processes in metaphor comprehension and memory. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 307–328). Cambridge, UK: Cambridge University Press.
- Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, *29*, 719–736.
- Wolff, P., & Gentner, D. (2011). Structure-mapping in metaphor comprehension. *Cognitive Science*, *35*(8), 1456–1488.
- Xu, X. (2010). Interpreting metaphorical statements. *Journal of Pragmatics*, *42*, 1622–1636.