Journal of Experimental Psychology: Learning, Memory, and Cognition

Forgetting the Literal: The Role of Inhibition in Metaphor Comprehension

Tim George and Jennifer Wiley


CITATION

RESEARCH REPORT

Forgetting the Literal: The Role of Inhibition in Metaphor Comprehension

Tim George and Jennifer Wiley
University of Illinois at Chicago

In order for a person to comprehend metaphoric expressions, do metaphor-irrelevant aspects of literal information need to be inhibited? Previous research using sentence-verification paradigms has found that literal associates take longer to process after reading metaphorical sentences; however, it is problematic to infer inhibition from this research. Moreover, previous work has not distinguished between familiar and novel metaphor processing. To test more directly for when inhibition may be required during metaphor processing, we performed 3 experiments using a metaphor-induced lexical forgetting paradigm. Participants initially learned word pairs where the cues were potential metaphoric vehicles and the targets were literal associates (e.g., SHARK–swim). Then, participants read half the vehicles as part of metaphorical sentences, which they interpreted (The lawyer for the defense is a shark). Subsequent forgetting of the literal associates was greater when vehicles had appeared in metaphorical sentences (Experiment 1) and was observed for both familiar and novel metaphors when participants were instructed to interpret the metaphors (Experiment 2) but was observed for only novel metaphors when participants were instructed to simply read the metaphors (Experiment 3). These results suggest that forgetting occurs as a result of inhibitory mechanisms that are engaged to alter activation of irrelevant literal information during metaphor processing, and that these mechanisms are most relevant for the processing demands associated with novel metaphors.

Keywords: metaphor, inhibition, retrieval-induced forgetting, figurative language

Supplemental materials: http://dx.doi.org/10.1037/xlm0000237.supp

Many theories have suggested that, compared with literal language, figurative language can involve additional processing (Bowdle & Gentner, 2005; Searle, 1979). On one hand, it has also been shown that people can sometimes understand metaphors quite directly (Glucksberg, Gildea, & Bookin, 1982), as in highly conventional or familiar metaphors (Blasko & Connine, 1993) where the meaning may already be stored in memory. On the other hand, it has been assumed that when one encounters a novel figurative phrase, processing using literal meanings will be attempted first, and nonliteral meanings may be considered only once a literal attempt has failed (Searle, 1979). When a novel metaphor is encountered, a person needs to engage in creative cognition to appreciate its meaning (Rutter et al., 2012; Silvia & Beaty, 2012).

For example, in the unconventional metaphor “Discipline is fertilizer,” the vehicle (fertilizer) is used to illustrate one or more aspects of the topic concept (discipline). The words discipline and fertilizer each have strong associates with other concepts but are weakly related as a pair. However, they share the semantic feature of promoting growth, which can be seen as a way to resolve the semantic anomaly presented by this sentence.

Theories of how metaphors are processed generally fall into two broad classes: comparison and categorization. Comparison accounts assume that metaphors are processed as similarity statements that highlight common properties between a topic and vehicle concept (Ortony, 1979) or that involve structural alignment between the two concepts, similar to analogy (Gentner, 1983; Gentner & Wolff, 1997; Wolff & Gentner, 2011). Categorization accounts suggest that the vehicle stands for a superordinate category of which the topic is a member (Glucksberg, McGlone, & Manfredi, 1997). However, the career of metaphor theory, an extension of the structure alignment view, suggests that these views are not mutually exclusive (Bowdle & Gentner, 2005). This theory suggests that unfamiliar metaphors are processed as comparisons but that as metaphors become conventionalized, there is a shift in processing toward categorization.

Each of these models involves excluding literal information from consideration, which suggests that inhibitory processes might be necessary to actively inhibit features of the vehicle that are not appropriate for the target. Previous research has indeed yielded findings consistent with inhibition of literal associates after read-
ing metaphorical sentences (Gernsbacher, Keysar, Robertson, & Werner, 2001; Glucksberg, Newsome, & Goldvarg, 2001). In these experiments, participants judged the validity of prime sentences containing both literal and metaphorical uses of a metaphor vehicle (e.g., That large hammerhead is a shark vs. That defense lawyer is a shark) before judging the validity of target sentences that were either relevant or irrelevant to the metaphorical meaning (e.g., Sharks are tenacious vs. Sharks are good swimmers). Participants were slower to verify metaphor-irrelevant properties following metaphorical sentence primes than following literal sentence primes, suggesting that inhibition is a mechanism by which irrelevant literal information is filtered out during metaphor processing.

However, because these studies relied on probe paradigms that used sentence verification times as the dependent measure, it is not clear whether these findings truly reflect the active suppression of metaphor-irrelevant information during processing. For example, in the Gernsbacher et al. (2001) study, it may be that once a prime is understood to contain a metaphorical meaning, it is more difficult to then integrate a conflicting literal-relevant target with the preceding meaning than when the literal target is preceded by a literal prime or other baseline prime. Such postreading probe measures make it difficult to conclude that inhibition must be recruited during the act of processing a metaphor. Additionally, these studies do not explicitly distinguish between familiar and novel metaphors, which is an important distinction, as indicated by the career of metaphor theory. If familiar metaphors simply point to a stored figurative representation, this suggests that their meaning can be accessed without the need to inhibit literal associates. However, for less-familiar metaphors that one has not previously encountered, literal information may initially interfere with resolving the metaphor. Consequently, one might expect novel metaphors to require more inhibition of literal information than familiar metaphors.

This distinction between familiar and novel metaphor processing is highlighted in a series of lexical decision experiments by Blasko and Connine (1993). They found similar response times for both literal and metaphorical target words immediately after reading familiar metaphorical sentences but longer response times for unrelated words, suggesting activation of both literal and figurative meanings versus unrelated control targets. This pattern was different for unfamiliar metaphors. For unfamiliar metaphorical sentences, there was evidence of activation of only the literal meaning immediately after reading. People were slower to respond to metaphorical targets than unrelated control targets at short stimulus onset asynchronies (SOAs). It was not until after a longer SOA that activation for metaphorical targets was found (Blasko & Connine, 1993). This pattern is consistent with the suggestion that familiar metaphors may have figurative meanings stored in memory that can be accessed along with the literal sense, whereas for novel metaphors figurative meanings may not be accessed until an initial literal-level processing attempt results in comprehension failure.

Thus, the goal of the present study was to test whether competition from and inhibition of literal associates during metaphor processing might be particularly important for novel metaphor expressions or whether they are involved for both familiar and novel metaphors. Instead of using response times on postreading recognition probes that are literal associates of a metaphorical vehicle, the current studies have adapted the retrieval-induced forgetting (RIF) framework (Anderson, Bjork & Bjork, 1994; Storm & Levy, 2012) to study metaphor processing. In RIF, as a consequence of attempting to retrieve target information in memory, subsequent recall of other related information is impaired, and this likely reflects inhibitory mechanisms. Analogously, in metaphor processing, any previously encountered irrelevant literal information that is activated upon presentation of a metaphor may have to be inhibited in order to effectively arrive at the figurative meaning, resulting in reduced recall for literal information. Based on prior work that has made a distinction in the processes recruited by familiar and novel metaphors, it is predicted that forgetting should be more likely to occur or will be stronger in its effect when metaphors are novel rather than familiar.

**Experiment 1**

Experiment 1 developed and tested a new paradigm for exploring processes underlying metaphor comprehension by extending the established RIF paradigm (Anderson et al., 1994; Storm & Levy, 2012). If inhibition of literal meaning occurs during the processing of metaphoric meaning, then recall for previously learned literal associates of metaphors should be reduced as a result of metaphor processing. In RIF experiments, participants typically study a list of category exemplar pairs (e.g., FRUIT–banana, FRUIT–lemon, TOOLS–hammer, TOOLS–saw), followed by a period of retrieval practice for half the exemplars from half the categories (e.g., FRUIT–banana). On a test of final recall for all of the word pairs, recall for responses associated with unpracticed categories (e.g., TOOLS) provides a baseline, and recall for practiced responses from practiced categories (e.g., FRUIT–banana) is enhanced. Critically, recall of unpracticed responses from practiced categories (e.g., FRUIT–lemon) is reduced (i.e., below baseline). This effect is assumed to be a consequence of competition arising from similar but irrelevant items in memory during retrieval practice, which in turn causes inhibition to be directed toward the competing information. Metaphor processing can be viewed as a form of selective retrieval from memory during which information competes for access. To the extent that competing or dominant responses such as literal associates are automatically brought to mind during metaphor processing, inhibitory mechanisms may be recruited to reduce these responses.

A variant of this paradigm called problem-solving-induced forgetting (PSIF) has already been extended into the domain of creative problem solving. PSIF has been demonstrated during the generation of solutions in creative tasks where subordinate meanings of words or unusual uses for objects are required, such as the remote associates test or the alternate uses test (Storm, Angello, & Bjork, 2011; Storm & Patel, 2014). These results are thought to demonstrate that inhibition is an adaptive mechanism that is engaged in order to deal with competition from less-creative solutions during creative problem-solving attempts.

In the context of metaphor processing, one would expect that the same mechanisms that underlie both RIF and PSIF would be recruited when a previously learned literal associate of a metaphor vehicle interferes with the successful comprehension of the metaphor’s figurative meaning. If readers engage in inhibition during the processing of metaphoric expressions, then participants should also show forgetting in this context, which would result in reduced recall for word pairs consisting of vehicles and their literal asso-
ciates. To test this hypothesis, we conducted an initial experiment in which participants studied metaphor-irrelevant literal associates of metaphor vehicles (e.g., SHARK–swim) before generating interpretations of metaphors that used half of these vehicles (e.g., The lawyer for the defense is a shark). This initial study relied on the materials of Gernsbacher et al. (2001) so as to demonstrate that an inhibitory effect could be replicated using this methodology. As in Gernsbacher et al., no distinction was made between novel and familiar metaphors in this first experiment. It was predicted that more forgetting would be observed for associates of the vehicles for which participants were shown and asked to interpret metaphoric expressions.

Method

Participants. Participants were 22 fluent English-speaking undergraduates from the University of Illinois at Chicago participant pool. All participants provided agreement to participate and in return received course credit.

Materials. The stimuli consisted of 40 metaphorical sentences taken from a previous study on metaphor comprehension (Gernsbacher et al., 2001). A set of 40 cue–response word pairs was created using these stimuli, in which the cue was the metaphor vehicle of the sentence and the response was a literal associate of the vehicle that did not relate to the figurative meaning (e.g., SHARK–swim for the sentence The lawyer for the defense is a shark). The single-word literal associates used in the word pairs for this study were based on the Gernsbacher et al. (2001) literal-meaning probes.

To assess base-rate recall of these word pairs, we conducted a norming experiment with 10 participants who studied the cue–response pairs on a sheet of paper for 3.5 m, followed by cue plus letter stem recall on a sheet of paper for 3.5 m. Average percentage recall was 95% (SD = 6%). For each individual item, the average percentage recall was calculated across participants. These stimuli were then split into two lists of 20 items that matched the target responses on base rate of recall, written frequency, number of letters, and number of syllables. The cue words were also matched for the number of living things. The two lists were used to counterbalance materials across the two conditions. This initial study relied on the materials of Gernsbacher et al. (2001) so as to demonstrate that an inhibitory effect could be replicated using this methodology. As in Gernsbacher et al., no distinction was made between novel and familiar metaphors in this first experiment. It was predicted that more forgetting would be observed for associates of the vehicles for which participants were shown and asked to interpret metaphoric expressions.

Participants. Participants were 22 fluent English-speaking undergraduates from the University of Illinois at Chicago participant pool. All participants provided agreement to participate and in return received course credit.

Materials. The stimuli consisted of 40 metaphorical sentences taken from a previous study on metaphor comprehension (Gernsbacher et al., 2001). A set of 40 cue–response word pairs was created using these stimuli, in which the cue was the metaphor vehicle of the sentence and the response was a literal associate of the vehicle that did not relate to the figurative meaning (e.g., SHARK–swim for the sentence The lawyer for the defense is a shark). The single-word literal associates used in the word pairs for this study were based on the Gernsbacher et al. (2001) literal-meaning probes.

To assess base-rate recall of these word pairs, we conducted a norming experiment with 10 participants who studied the cue–response pairs on a sheet of paper for 3.5 m, followed by cue plus letter stem recall on a sheet of paper for 3.5 m. Average percentage recall was 95% (SD = 6%). For each individual item, the average percentage recall was calculated across participants. These stimuli were then split into two lists of 20 items that matched the target responses on base rate of recall, written frequency, number of letters, and number of syllables. The cue words were also matched for the number of living things. The two lists were used to counterbalance materials across the two conditions, with equal numbers of participants in each condition.

Procedure. After agreement to participate was collected, there were four phases to the main experiment: study, initial recall, metaphor interpretation, and final recall. First participants studied the word pairs. They were given the 40 cue–response word pairs on a sheet of paper and instructed to study each word pair so that when given the cue word, they could recall the response word. The order of items was randomized for each participant. They had 3.5 m to study this list before it was collected.

Immediately following the study period, participants were given a list of the 40 cue words along with the first letter of the response word on a sheet of paper in a new random order. They had 3.5 m to write down the correct response to each cue before it was collected. This procedure ensured strengthening of the cue–response pairs in memory.

Following the initial recall, the metaphor interpretation task was presented electronically using the E-Prime 2.0 software (Psychology Software Tools Inc., 2012). Participants were presented with one of the two lists of 20 metaphorical sentences in which the vehicle was a cue from the previously studied word pairs (metaphor condition). The other half of the cue words did not appear in any sentences (no-metaphor condition). The sentences were presented in random order one at a time. For each sentence, participants had up to 20 s to read and think of the metaphorical interpretation, which they indicated by pressing the space bar. The computer logged the response time for this space bar press. After pressing the space bar, they had up to 20 s to type the interpretation in a response box before automatically being presented with the next sentence.

Following the metaphor interpretation task, participants received a surprise final recall test of all 40 cue–response pairs via E-Prime. The order of items was randomized across participants. Each trial began with the presentation of a fixation cross for 500 ms, followed by the presentation of the cue word along with a blank response box for 5 s. Participants were instructed to type the correct response into the response box. After 5 s, the next trial began.

Results

The pattern of final recall is shown in Table 1. The percentage of correct responses on the final recall test was computed separately for the metaphor interpretation and no-metaphor interpretation conditions to create recall accuracy scores for each condition. Misspellings and pluralizations were accepted as correct. These values were entered into a repeated-measures analysis of variance (ANOVA). More forgetting of the literal associates was seen in the metaphor-interpretation condition than in the no-metaphor condition, $F(1, 21) = 12.05, p < .01, \eta^2_p = .37; F(1, 39) = 20.74, p < .01, \eta^2_p = .35$, suggesting metaphor-induced forgetting. There were no differences in initial recall for either participant or item analyses ($F_1 < 1; F_2 < 1; M = 91\%$, $SD = 9\%$).

Discussion

Recall of literal associates of metaphor vehicle words was reduced following metaphor interpretation, compared to when those words did not appear in metaphors. The finding that there is forgetting of literal associates of metaphoric vehicles as a result of generating their figurative interpretations is consistent with the notion that inhibition of irrelevant literal information is involved in metaphor processing (Gernsbacher et al., 2001; Glucksberg et al., 2001).

Experiment 2

The results of Experiment 1 suggest that inhibition of irrelevant literal associations may occur as part of metaphor processing. However, a great deal of research has suggested that there may be

<table>
<thead>
<tr>
<th>Condition</th>
<th>$M$</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metaphor</td>
<td>76%</td>
<td>19%</td>
</tr>
<tr>
<td>No-metaphor</td>
<td>85%</td>
<td>16%</td>
</tr>
</tbody>
</table>
an important distinction between the processing of familiar and novel metaphors (Blasko & Connine, 1993; Bowdle & Gentner, 2005). Familiar metaphors may be processed more like categorical statements (e.g., “A bear is a mammal”), whereas novel metaphors may require an initial comparison process in order to arrive at an appropriate metaphorical meaning.

The Gernsbacher et al. (2001) and Glucksberg et al. (2001) studies do not make such a distinction. A novel metaphor vehicle is not as likely to have a stored figurative representation or associated metaphorical category as is a familiar metaphor vehicle. This should make processing more challenging and may require more inhibitory processing in order to overcome the irrelevant literal meaning. This may result in even more forgetting of literal associates for novel metaphors. To test this hypothesis, for this experiment we had participants study literal associates of both familiar and novel metaphor vehicles before generating interpretations of metaphors that used half of these vehicles within each type (familiar and novel). If readers engage in inhibition particularly during the processing of novel metaphor expressions, it was predicted that more forgetting would be observed for associates of the vehicles for which participants were asked to interpret novel metaphor expressions than for familiar metaphor expressions.

Method

Participants. Thirty fluent English-speaking undergraduates from the University of Illinois at Chicago participant pool provided agreement to participate and were given course credit for their participation.

Materials. The original 40 metaphorical sentences from Gernsbacher et al. (2001) stimuli along with 30 new metaphorical sentences were rated on familiarity by seven raters on a 7-point scale ranging from 1 (not all familiar) to 7 (highly familiar). From this sample we selected 20 metaphors with the lowest familiarity ratings (M = 2.23, SD = 0.62) to serve as stimuli for the novel metaphor condition (e.g., My mother says envy is rust) and 20 metaphors with the highest familiarity ratings (M = 5.89, SD = 0.64) to serve as stimuli for the familiar metaphor condition (e.g., Her husband is a gem). This difference in ratings was significant t(38) = 18.40, p < .01. Additionally, latent semantic analysis was used to provide a measure of semantic similarity between each metaphor’s vehicle and topic terms. Similarity was higher for the familiar metaphors (M = .12, SD = .10) compared with the novel metaphors (M = .04, SD = .05), t(38) = 2.23, p < .01. The sentences used similar syntactic structures; were similar in word count between the familiar (M = 6.0, SD = 1.3) and novel conditions (M = 6.6, SD = 1.2), t(38) = -1.57, p > .05; and similar in syllable count between the familiar (M = 9.0, SD = 2.7) and novel conditions (M = 10.2, SD = 2.2), t(38) = -1.59, p > .05. Adjusting response time measures for sentence length did not change the pattern of results for either Experiment 2 or Experiment 3.

A set of 40 cue–response word pairs were created using these stimuli, in which the cue was the metaphor vehicle of the sentence and the response was a literal associate of the vehicle unrelated to the figurative meaning (e.g., RUST–red). Base-rate recall for the target responses was tested in a norming study with 10 participants. It followed the same base-rate assessment procedure as in the norming study used for Experiment 1, with the addition of a second recall phase completed via E-prime which was identical to the final recall procedure of Experiment 1. This ensured that words appearing in the different conditions were matched according to recall rate on the second test without any intervening task. Average percentage recall was 94% (SD = 10%) for the familiar vehicles and 94% (SD = 9%) for the novel vehicles. The stimuli were split into two lists of 20 items, each list containing 10 items associated with novel metaphors and 10 items associated with familiar metaphors. The two lists were used to counterbalance materials across the two conditions, with equal numbers of participants in each condition. Each of the four subsets of 10 items was matched for rate of recall on the second test, written frequency, number of letters, and number of syllables. The cue words were matched on the number of living things.

Procedure. The procedure and scoring of the final recall test was identical to that of Experiment 1.

Results

Analysis of response times revealed that participants spent longer thinking of interpretations of novel metaphors (M = 7.6 s, SD = 2.6) than familiar metaphors (M = 4.9 s, SD = 1.6), F(1, 29) = 72.26, p < .01, ηp2 = .71, suggesting greater processing difficulty for novel metaphors.

An examination of average initial recall for all items revealed one outlier that was more than 3 SDs below the mean, so this item was dropped from analyses. The percentage of correct responses on the initial and final recall tests were computed separately for the familiar and novel metaphor conditions and the two no-metaphor conditions to create four recall accuracy scores for each participant for both initial and final tests. When the initial recall scores (M = 97%, SD = 4%) were entered into a two-way repeated measures ANOVA, no differences were seen due to either metaphor condition (F1 < 1; F2 < 1) or familiarity (F1 < 1; F2 < 1).

As shown in Table 2, when the final recall scores were entered into a two-way repeated-measures ANOVA, there was a main effect of metaphor condition, F1(1, 29) = 11.33, p < .05, ηp2 = .28; F2(1, 37) = 7.53, p < .05, ηp2 = .17, such that recall for items appearing in metaphors was lower than recall for items not appearing in metaphors. There was also a main effect of familiarity, F1(1, 29) = 13.67, p < .05, ηp2 = .32; F2(1, 37) = 4.51, p < .05, ηp2 = .11, such that recall for response words associated with novel vehicles was lower than for words associated with familiar vehicles, regardless of whether they appeared in metaphor sentences. The interaction was not significant (F1 < 1; F2 < 1).

Discussion

The results of Experiment 2 replicate the metaphor-induced forgetting effect seen in Experiment 1. When participants read and were prompted to generate an interpretation for a metaphor expression, they were less able to retrieve the associate they had learned previously for the vehicle. The results also indicated a
main effect for familiarity condition such that the words associated with the novel vehicles were recalled less at final recall regardless of whether they appeared in metaphor sentences. This result was unexpected because the norming study had suggested that the recall rates would be more similar in the no-metaphor condition and because recall rates were matched at initial recall. Further, contrary to the strong prediction that more forgetting would be seen for novel expressions, a forgetting effect was found for even familiar metaphors.

One explanation for this general forgetting effect might be the methodology that was used in the first two studies. Requiring participants to explicitly type in an interpretation of a familiar metaphor might serve as a source of interference that blocks access to the initially learned associates. In contrast to forgetting via an inhibitory route, the blocking account serves as a common alternative explanation for many RIF results (MacLeod, Dodd, Sheard, Wilson & Bibi, 2003). In an attempt to reduce the likelihood of blocking or interference as a source of forgetting, in Experiment 3 participants were asked to only read the metaphoric statements and not interpret them. Removing the instruction to generate an explicit interpretation was intended to eliminate a potential source of interference, thereby providing a closer measure of inhibitory-based forgetting as a result of normal reading processes.

**Experiment 3**

Because instructing participants to produce a written interpretation for each metaphor may have been partly responsible for the forgetting effects seen in the previous experiments, Experiment 3 tested for forgetting effects when participants were instructed to simply read metaphors in a reading-only condition.

Instead of generating written interpretations of the metaphoric sentences, participants in the reading-only condition were told to press a key when they were done reading the sentences. This condition allowed us to see metaphor-induced lexical forgetting patterns in a more-natural reading context (without the explicit requirement of typing in an interpretation). If inhibition is utilized during the course of simply reading novel metaphors, then forgetting of literal associates of novel metaphor vehicles should be still observed, as it was in Experiment 2. However, if inhibition is not as necessary during the course of simply reading familiar metaphors, then no such forgetting should be observed, in contrast to the findings of Experiment 2. This would suggest that the forgetting that did result from interpreting familiar metaphors in Experiment 2 was due to interference arising from participants’ explicitly generated interpretation.

### Table 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Familiar</th>
<th>Novel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Metaphor</td>
<td>86%</td>
<td>13%</td>
</tr>
<tr>
<td>No-metaphor</td>
<td>92%</td>
<td>10%</td>
</tr>
</tbody>
</table>

**Method**

**Participants.** Participants consisted of 26 fluent English-speaking undergraduates from the University of Illinois at Chicago participant pool who provided agreement to participate and in return were given course credit.

**Materials.** The same materials as in Experiment 2 were used.

**Procedure.** The procedure was similar to that of Experiment 2 except that participants were instructed to simply carefully read each sentence and to press a key after doing so.

### Results

Examination of reading times revealed that one participant’s reading time (18.1 s) was over 3 SDs above the mean. After excluding this participant, participants took 3.9 s (SD = 1.6) to read familiar metaphors and 4.7 s (SD = 1.7) to read the novel metaphors, \( F(1, 25) = 24.22, p < .01, \eta^2_p = .50 \), indicating that metaphor novelty still impacted processing time, despite the lack of a requirement to generate an interpretation. Initial recall scores were entered into an ANOVA crossing metaphor condition and familiarity. There were no effects of metaphor condition \( (F_1 < 1; F_2 < 1) \), familiarity \( F(1, 24) = 2.41, p = .13, \eta^2_p = .09; F(1, 38) = 1.90, p = .18, \eta^2_p = .05, \) or an interaction, \( F(1, 24) = 1.70, p = .20, \eta^2_p = .06; F(1, 38) = 2.26, p = .14, \eta^2_p = .06 \).

For the final recall data, the ANOVA revealed a main effect of metaphor condition, \( F(1, 24) = 7.19, p < .05, \eta^2_p = .23; F(1, 38) = 3.45, p = .07, \eta^2_p = .08, \) with more forgetting in the metaphor condition than in the no-metaphor condition and a marginal effect of familiarity, \( F(1, 24) = 3.69, p = .07, \eta^2_p = .13; F(1, 38) = 0.83, p = .38, \eta^2_p = .02, \) with greater forgetting associated with novel than familiar metaphors. The interaction of metaphor condition and novelty was also significant, \( F(1, 24) = 4.81, p < .05, \eta^2_p = .17; F(1, 38) = 4.58, p < .05, \eta^2_p = .11. \) As shown in Table 3, there was more forgetting of literal associates from having read novel metaphors compared with the no-metaphor condition, \( F(1, 24) = 10.00, p < .01, \eta^2_p = .29; F(1, 19) = 6.12, p < .05, \eta^2_p = .24, \) but no forgetting was observed for reading familiar metaphors compared with no-metaphors \( (F_1 < 1; F_2 < 1) \).

### Discussion

In Experiment 3, forgetting was still observed under reading-only instructions, which supports the hypothesis that forgetting did not depend on forcing participants to generate a written response. Critically, the interaction between familiarity and metaphor exposure was significant, because the forgetting effect was observed for only novel metaphors and not for familiar metaphors.

### Table 3

<table>
<thead>
<tr>
<th>Condition</th>
<th>Familiar</th>
<th>Novel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Metaphor</td>
<td>84%</td>
<td>16%</td>
</tr>
<tr>
<td>No-metaphor</td>
<td>83%</td>
<td>14%</td>
</tr>
</tbody>
</table>
General Discussion

Across all three experiments, following the processing of metaphoric sentences, recall of literal associates of metaphor vehicles that were irrelevant to the metaphoric meaning of the sentences was impaired. However, differences in patterns between Experiments 2 and 3 suggest that there may be two different sources of forgetting in these studies. In Experiment 2, forgetting may have resulted from postcomprehension interference from the metaphor interpretation, which prevented access to the initially learned item. In Experiment 3, a different pattern emerged. Forgetting was observed under reading-only instructions, in which no such explicit interpretation had to be articulated by the reader. Moreover, this forgetting effect was observed for only novel metaphors and not for familiar metaphors. If forgetting resulted purely from interference arising from the metaphoric meaning or from the formation of a new associative relationship in the metaphor processing task, one would not expect metaphor novelty to differentially affect this effect. In fact, one might argue that forgetting would be stronger for familiar metaphors, due to their highly salient meaning. Thus, although an interference account may provide the best fit for the pattern of results seen in Experiment 2, it cannot explain the pattern seen in Experiment 3.

This pattern instead suggests that the locus of the forgetting effect lies earlier in the comprehension process and is a result of inhibitory mechanisms. Specifically, we suggest that upon encountering a novel metaphor vehicle in a sentence, a person’s attempts at integrating the vehicle and topic are unsuccessful due to inappropriate meaning activation. This failure then drives a selection process that requires inhibition of the episodic trace containing the previously learned literal associate, which results in forgetting. On the other hand, upon encountering a familiar metaphor vehicle, one’s attempts at integrating the vehicle and topic may be more successful if the vehicle readily activates a known metaphorical category. This reduces the need for inhibition, and therefore less forgetting is seen.

This explanation is compatible with both categorization (Glucksberg, 2008) and comparison (Bowdle & Gentner, 2005) models of metaphor processing. These models suggest that, generally, metaphors are processed as categorization statements when possible. As described earlier, for familiar metaphors, it may be relatively easy to select the relevant properties from the irrelevant properties because the vehicle word strongly activates a superordinate category of meaning. Because there is no competition during retrieval of the appropriate meaning, category-based processing does not result in forgetting. However, attempts at processing novel metaphors via such a categorization process will fail. This will prompt the reader to engage in a selection process that involves activating many features of the topic and vehicle in order to compare them and find an alignment. Although novel metaphors might also be understood as categorizations once they are processed and understood, achieving this resolution requires a search process in which irrelevant features of the vehicle are likely to be activated and thus may need to be inhibited, resulting in forgetting.

These results suggesting inhibitory processes are also consistent with individual difference findings implicating the need for control mechanisms to inhibit irrelevant information during metaphor processing (Chiappe & Chiappe, 2007) and modeling work that requires inhibitory connections to features of the metaphor vehicle that are less important (Kintsch, 2000). Additionally, these findings dovetail with neuroimaging research indicating increased activity in left inferior frontal gyrus while reading novel metaphoric or nonmeaningful sentences compared with literal sentences (Rutter et al., 2012; Stringaris, Medford, Giampietro, Brammer, & David, 2007). Activation in this area has been associated with controlled semantic retrieval (Wagner, Paré-Blagoev, Clark & Poldrack, 2001) as well as with RIF (Wimber et al., 2008).

There are some limitations to this work that should be noted. For instance, this work compared forgetting effects for familiar and novel metaphor stimulus sets that could not be matched completely (i.e., it was not possible to change metaphors from familiar to novel by using the same sentence frame and changing only a single word). Thus, there is a need to replicate these effects using better matched literal control sentences as comparisons. In addition, it seems important to add comprehension checks for the read-only conditions. More work is also needed to address how other properties of metaphors, such as aptness and topic constraint, might affect forgetting patterns. It would also be worthwhile to extend this work to other linguistic tropes, including idioms, sarcasm, and irony.

Conclusion

In summary, the results of these experiments extend previous work showing forgetting as a consequence of creative problem solving in which irrelevant information must be overcome (Storm et al., 2011; Storm & Patel, 2014). These findings indicate that forgetting of irrelevant literal information may be a useful part of novel figurative language processing. When one encounters such creative uses of language, successful comprehension may partly depend on reducing the activation of irrelevant literal information in memory. Particularly the results of the third experiment suggest this is likely a reflection of inhibitory processes that are utilized to aid processing.

References
