

## Evidence for Role-Neutral Initial Processing of Metaphors

Phillip Wolff  
University of Maryland

Dedre Gentner  
Northwestern University

Two models of metaphor processing are contrasted. The structure-mapping model postulates an initially role-neutral alignment process, followed by directional projection of inferences. The attributive categorization model postulates role-specific processing throughout comprehension. To test between these models, the early stages of metaphor comprehension were probed using a technique based on S. Glucksberg, P. Gildea, and H. Bookin's (1982) finding that metaphorical meaning interferes with literal truthfulness judgments. In Experiment 1, interference effects did not differ between normal metaphors and metaphors with reversed terms, suggesting that initial processing is role-neutral. In Experiment 2, we again found no role dependence in interference effects, even for highly conventional metaphors. In Experiment 3, it was verified that (a) full comprehension is role-sensitive and (b) full comprehension reaction times (RTs) are far longer than interference RTs, buttressing the claim that interference is an early-stage effect. Overall, the results support the structure-mapping model of metaphor processing.

Metaphors both highlight commonalities and invite new inferences. Whereas finding commonalities can be conceived of as a symmetrical matching process, the inference projection process is by nature asymmetrical; information is directionally projected from the vehicle (or *base*) to the topic (*target*). For example, in a nominal metaphor such as *A rumor is a virus*, ideas such as contagion that are normally associated with the vehicle concept, *virus*, are projected to the topic concept, *rumor*. The role specificity of metaphoric processing is revealed when the ordering of the terms is changed (Glucksberg & Keysar, 1990; Miller, 1993; Ortony, 1979; Ortony, Vondruska, Foss, & Jones, 1985). For example, if the terms of the aforementioned metaphor are reversed, the resulting statement, *A virus is a rumor*, seems pointless. In other cases, reversing the terms produces a change in meaning. For example, the metaphor *My surgeon is a butcher* suggests cutting sloppily, whereas its reversal, *My butcher is a surgeon*, suggests cutting precisely. Models of metaphor and analogy differ in their assumptions as to when in processing this asymmetry appears. The key dichotomy here is between models that explain metaphors in terms of comparison processing and models that explain metaphors in terms of category processing.

Phillip Wolff, Department of Psychology, University of Maryland; Dedre Gentner, Department of Psychology, Northwestern University.

This research was supported by National Science Foundation Grant SBR-9511757 and by Office of Naval Research Grant N00014-92-J-1098 awarded to Dedre Gentner. We thank Brian Bowdle, Sarah Brem, Evan Heit, Arthur Markman, and the entire Similarity and Analogy Group at Northwestern University for many discussions of these issues. We also thank Sam Glucksberg and two anonymous reviewers for their suggestions.

Correspondence concerning this article should be addressed to Phillip Wolff, Department of Psychology, University of Maryland, Zoology–Psychology Building, College Park, Maryland 20742. Electronic mail may be sent to wolff@psyc.umd.edu.

### Two Models of Metaphor Processing

One way to model metaphor is to consider it a kind of analogical comparison. According to Gentner's structure-mapping model, role sensitivity arises late in the time course of processing (Falkenhainer, Forbus, & Gentner, 1989; Forbus, Gentner, & Law, 1995; Gentner & Markman, 1997; Gentner & Wolff, 1997; Wolff & Gentner, 1992). The initial stage of metaphor processing is symmetrical—that is, role neutral. Elements of the two representations are placed in alignment on the basis of local identities and common roles in the larger systems of knowledge (Clement & Gentner, 1991; Markman & Gentner, 1993b). Once a common system is in place, role information is used to project directional inferences. Predicates connected to the common structure in the base, but not initially present in the target, are imported into the target as candidate inferences. For example, given the metaphor *My surgeon is a butcher*, an initial (symmetrical) alignment process would yield the common system “one who cuts flesh.” Then, role-specific inference processes project further ideas from the base to the target (e.g., the idea of cutting crudely and without regard for the health of the flesh). Thus, in structure-mapping, processing begins symmetrical and ends directional.

Another major class of models considers metaphor as akin to category inclusion and assumes that processing is directional from the beginning (Glucksberg & Keysar, 1990, 1993). In Glucksberg, McGlone, and Manfredi's (1997) attributive categorization model, potential categories are generated and projected from the vehicle during metaphor processing, while sets of modifiable dimensions are simultaneously identified in the topic. For example, in the aforementioned metaphor, a category like “one who cuts flesh crudely” is derived from the vehicle, *butcher*. Meanwhile, the topic, *surgeon*, yields dimensions that can be modified: degree of precision, income, level of activity, and so on. The interpretation of the metaphor is thus a kind of negotiation

between the category prototypically associated with the vehicle and the dimensions of description of the topic. As Glucksberg et al. stated,

Instead of property matching, informative comparisons must be understood via a property attribution process. At a minimum, this process must involve the selection of one or more candidate properties from the vehicle concept, and an attempt to apply them to the topic concept. (p. 51)

Thus, Glucksberg et al. made clear that the vehicle and topic are processed in role-sensitive ways from the very beginning of comprehension.

In both the structure-mapping and the attributive categorization models, role information is crucial in the comprehension process. The key difference concerns when in the time course of comprehension this information is used. According to the structure-mapping model, processing is initially symmetrical, or role neutral; not until the stage of inference projection does processing become role specific (asymmetrical). According to the attributive categorization model, processing is role specific from the beginning.

We tested whether early processing is role specific by varying the order of terms, contrasting *forward metaphors* (e.g., *A rumor is a virus*) with *reversed metaphors* (e.g., *A virus is a rumor*). Both theories predict that these should be processed differently, but structure-mapping theory predicts early equivalence, whereas property attribution theory predicts differential processing throughout. In Experiments 1 and 2, we examined the initial stages of comprehension, using a metaphoric interference task adapted from Glucksberg, Gildea, and Bookin's (1982) studies. In Experiment 3, we investigated the later time course of comprehension. We first describe the metaphoric interference task.

### The Metaphoric Interference Technique

To investigate early processes in metaphor comprehension, we used a technique developed by Glucksberg et al. (1982) in their landmark studies and extended by Keysar (1989). In Glucksberg et al.'s task, participants were shown three basic kinds of statements: true class-inclusion statements (e.g., *Some birds are robins*), false (anomalous) class-inclusion statements (e.g., *Some birds are apples*), and metaphorical statements (e.g., *Some jobs are jails*). Participants were asked to classify these statements as either literally true or literally false. As expected, participants could speedily classify class-inclusion statements as literally true and anomalous statements as literally false. However, participants had difficulty classifying metaphors; they were slower to respond "literally false" to metaphors than to ordinary false (i.e., anomalous) statements. The fact that metaphoric meaning interfered with literal true-false judgments implies that processing of the metaphors' meanings began before the entire literal meaning was processed. As Glucksberg et al. (1982) suggested, "people seem to process both the nonliteral and literal meanings of sentences in the same way, and at the same time" (p. 85). These results were a decisive strike against two-process accounts of metaphor comprehension, according to which metaphors are first

processed as literal meanings, and only if literal processes fail are metaphoric processes invoked.

Glucksberg et al.'s (1982) metaphor interference effect is strong evidence for early processing of metaphorical meaning. For our present purposes, the beauty of this technique is that it appears to tap into the initial processing stages; the interference effect does not require that processing be completed, only that it be initiated early enough to interfere with a literal true-false judgment. (We verify this intuition in Experiment 3; full metaphor comprehension requires considerably more time than the time required to reject these same metaphors in the interference task of Experiments 1 and 2.) The metaphor interference technique, in conjunction with a manipulation of role information, allows a test of the two models.

According to the attributive categorization model, initial processing is role specific. If the terms are in reversed order (e.g., *A warehouse is a brain*), the sentence should simply seem anomalous, because there is no sensible projection of a category from the vehicle to the dimensions provided by the topic. Thus, metaphoric interference is expected only when the terms are in the forward order (*A brain is a warehouse*). A reversed ordering of the terms should lead to no more interference than would an anomalous statement.

In contrast, if metaphors are processed as figurative comparisons, then the initial process should be a role-neutral process of structural alignment. Hence, any metaphoric interference effects should be independent of the order of the terms.

### Testing the Directionality of Metaphors

The logic of this research requires that the metaphors used be clearly directional. The test is unfair to the attributive categorization model if the metaphors lack a clear forward direction (for in such a case, the forward and reversed metaphors would differ only in the experimenters' labels). Only if the metaphors have a clearly preferred direction of interpretation is the category projection postulated by the attributive categorization model predicted to be more natural in the forward direction than in the reversed direction. To put it another way, the finding that processing is initially role neutral is interesting only if the metaphors used are clearly *not* role neutral in later stages of processing. Therefore, we needed highly directional metaphors; that is, we needed metaphors that become anomalous when reversed (e.g., *Some jails are jobs*). Many of the metaphors used in Glucksberg et al.'s (1982) studies could not be used in this experiment because their reversals resulted in interpretable metaphors (e.g., *Some surgeons are butchers*, *Some butchers are surgeons*).

To ensure high directionality, we carried out two directionality rating tasks. First, directionality ratings for 50 metaphors were collected from 25 Northwestern undergraduates. Each metaphor was shown in forward and reversed ordering together on a page, and participants rated each ordering on a 1 to 5 scale of goodness. The 30 metaphors with the largest advantage for forward direction over reversed direction were selected for use.

As a further check, we obtained directionality ratings for these 30 metaphors from 16 Northwestern undergraduates. Participants saw the forward and reversed versions of each metaphor simultaneously on a computer screen and chose which version they preferred. Each participant saw all 30 metaphors; presentation order was randomized across participants. Position of the two orderings on the screen, one on top of the other, was counterbalanced across four between-subjects groups. The forward versions of the metaphors were chosen 89% of the time.

### Experiment 1

The goal of Experiment 1 was to discover whether the initial processes used in metaphor comprehension are role neutral (symmetrical), as predicted by structural alignment theory, or asymmetrical, as predicted by property attribution theory. Following Glucksberg et al. (1982) and Keysar (1989), we embedded the metaphors in a literal true-false task. The dependent measure was the participants' speed at classifying statements as either literally true or literally false. There were four kinds of literally false statements: (a) ordinary false (anomalous) statements (e.g., *Some birds are apples*), (b) forward metaphors (e.g., *Some jobs are jails*, *Some rumors are viruses*), (c) scrambled metaphors (e.g., *Some rumors are jails*, *Some jobs are viruses*), and (d) reversed metaphors (e.g., *Some jails are jobs*, *Some viruses are rumors*). There were also two kinds of literally true statements: high-typicality statements, such as *Some birds are robins*, and low-typicality statements, such as *Some birds are penguins*. These two kinds of literally true statements were included as manipulation checks. If participants process sentences as they normally do in sentence verification experiments, they should be faster to classify high-typicality than low-typicality statements as literally true (e.g., McCloskey & Glucksberg, 1979; Smith, Shoben, & Rips, 1974).

The key predictions concern the false statements. The first question was whether Glucksberg et al.'s (1982) results would be replicated—that is, whether (forward) metaphors would take longer to classify as false than would anomalies or scrambled metaphors, indicating that metaphorical meanings are processed early, before finishing literal processing. If so, then the stage is set for the key differentiating prediction. If initial processing is role neutral, as predicted by structural alignment theory, then an interference effect should be found for reversed metaphors as well as for forward metaphors. If initial processing is directional and role specific, as predicted by the attributive categorization model, forward metaphors should have produced more interference than reversed metaphors (which should have produced little or no interference).

### Method

**Participants.** The participants were 36 Northwestern undergraduates who participated for course credit. The data of two participants were discarded because of excessive error rates (over 30%).

**Materials.** Two 180-item test lists were constructed, each with 90 literally true statements (45 high, 45 low typical), 30 anomalies, 30 metaphors, and 30 scrambled metaphors. The metaphors that were used are shown in Appendix A. In any given list, the terms for half of these metaphors were in the forward direction and half were in the reversed direction. The test items were preceded by 12 warm-up (filler) items, 6 true and 6 anomalous statements. The literally true and the anomalous statements were the same as those used by Glucksberg et al. (1982). For the literally true statements, typicality was based on high- and low-frequency items from each of 15 categories of the Battig and Montague (1969) category-exemplar production norms. Anomalous statements were constructed by scrambling the subject and predicate nouns of the literally true statements.

The metaphors used were derived from a larger set that included many of Glucksberg et al.'s (1982) original metaphors. As described in *Testing the Directionality of Metaphors*, all of the metaphors were rated as highly directional. Reversed metaphors were constructed by reversing the topic and vehicle terms of the forward metaphors. In addition, scrambled metaphors were included to ensure that any metaphor effects could not be attributed to the particular words that appear in the metaphors. These were constructed as in Glucksberg et al.'s research, by re-pairing the subject and predicate nouns of the metaphors to produce inappropriate pairings.

In addition, a 120-item practice list (60 false statements and 60 true statements) was constructed. This was similar to the practice list used by Glucksberg et al. (1982) except that Glucksberg et al. used 15 forward metaphors and 15 scrambled metaphors, whereas we used 10 metaphors, 10 scrambled metaphors, and 10 reversed metaphors. In addition, 30 anomalous statements and 60 literally true statements (30 high typical and 30 low typical) were taken from Glucksberg et al. Both the practice and the test lists were individually randomized for each participant.

**Procedure.** Participants used DOS-based computers separated by sound attenuating carrels. Participants were told that they would see statements like *Some zebras are camels* and that they should say whether these statements were literally true or false by pressing the left or right arrow keys, respectively. Participants were then given 20 practice trials pressing the left and right arrow keys in response to the phrases "Literally true" or "Literally false," respectively. Errors were indicated. The participants then received 120 practice trials on statements like those in the test phase. Presentation of the sentences in both the practice and the test sessions occurred in the following manner: First, a line of pound signs appeared to the right of a > symbol. The number of pound signs matched the number of letters in the sentence. After 300 ms, the line was replaced with a sentence, which remained on the screen until an arrow key was pressed. The dependent measure was the time between the appearance of the sentence and the pressing of an arrow key. In both the practice and the test sessions, if participants failed to respond within 5,000 ms of the sentence's appearance, a soft bell sounded, accompanied by a message saying that the response was too slow. After the practice sessions, a short message appeared, reminding participants to respond as quickly and as accurately as possible. In all of the practice and test sessions, incorrect responses were pointed out. Sentence order was randomized for each participant. We achieved millisecond accuracy by using Pascal procedures provided in Brysbaert, Bovens, and d'Ydewalle (1989).

**Design.** A within-subject design was used with each participant verifying each sentence type. Metaphor direction (forward vs. reversed) was counterbalanced across two between-subjects groups. Order of items was randomized across participants.

## Results

Incorrect responses (7%) and reaction times (RTs) greater than 2,400 ms (2%) were removed, resulting in the exclusion of 9% of the data. The results show early symmetric processing, as predicted by structure-mapping theory. As shown in Figure 1, reversed metaphors produced as much interference as forward metaphors. These observations are supported by statistical analyses using subjects ( $F_s$ ) and items ( $F_i$ ) as the random variables. Focusing on the four different kinds of false statements, a one-way repeated measures analysis of variance (ANOVA) indicated an overall effect of sentence type,  $F_s(3, 99) = 4.43, p < .001, MSE = 4,973; F_i(3, 86) = 2.72, p < .05, MSE = 7,470$ . To examine specific predictions, we conducted individual comparisons, based on a pooled error term, using a .0125 significance level derived from the Bonferroni method of error correction.

As expected, forward metaphors ( $M = 1,118$  ms,  $SD = 204$ ) took longer to judge as false than did scrambled metaphors ( $M = 1,078$  ms,  $SD = 169$ ),  $t_s(33) = 3.31, p < .01; t_i(58) = 3.01, p = .01$ . The results thus replicate the main finding of Glucksberg et al. (1982). Of primary interest, however, is that reversed metaphors ( $M = 1,111$  ms,  $SD = 194$ ) took longer to judge as false than did scrambled metaphors,  $t_s(33) = 2.73, p < .01; t_i(58) = 2.84, p < .01$ . Thus, contrary to the attributive categorization model, interference effects did not disappear when the canonical order of the terms was reversed.

More important, there was no indication that the degree of interference produced by reversed metaphors was any less than that produced by forward metaphors,  $t_s(33) = 0.58, ns; t_i(58) = 0.16, ns$ . These similar levels of interference cast doubt on the possibility that reversed metaphors were identified as peculiar and then quickly reversed to produce interference in a task that had nothing to do with metaphors. As expected, RTs for scrambled metaphors did not differ from those for anomalous statements ( $M = 1,065$  ms,  $SD = 167$ ),  $t_s(33) = 1.08, ns; t_i(58) = 2.00, ns$ , which suggests that the scrambled metaphors were nonmeaningful,

as intended. Finally, the error rates for literally true statements (8%), anomalous statements (9%), scrambled metaphors (2%), forward metaphors (4%), and reversed metaphors (5%) were moderately low, indicating that participants' concentration levels during the experiment were reasonably high.

## Discussion

According to structural alignment theory, the initial processing of comparisons is role neutral. The key result supporting this claim is that reversed metaphors, like forward metaphors, took longer to reject than did scrambled metaphors. Indeed, the degree of interference produced by reversed metaphors did not differ from that of forward metaphors. The near equivalent levels of interference suggest similar early processing and render improbable the possibility that the reversed metaphors were processed by switching the terms.

Three points emerge from this study. First, the metaphoric interference effect shows that processing of the metaphoric sentences was initiated very early, before the literal judgment was completed, which is consistent with the Glucksberg et al. (1982) account. Second, the symmetrical nature of the interference effect shows that this early processing was initiated on the basis of the terms' meanings, independently of their vehicle–topic (base–target) roles in the metaphor. Third, the fact that both forward and reversed metaphors created interference, but scrambled metaphors (which used the same words) did not, shows that this early processing is sensitive to interactions between the two meanings. (It also emphasizes that the key result is not the null effect of no difference between the forward and reversed metaphors; rather, the main result is the *difference* in interference between *both* forward and reversed metaphors and scrambled metaphors.) This pattern is as predicted by the early alignment process postulated by the structure-mapping theory.

These findings are inconsistent with the prediction of initially asymmetrical, role-specific processing made by Glucksberg, McGlone, and Manfredi's (1997) attributive categorization model. If the vehicle and the topic of a metaphor contribute differently to the metaphor's interpretation from the beginning, then the reversed metaphors, which do not afford appropriate interaction of potential categories and potential dimensions, should not create interference. Consider the example metaphor *Some salesmen are bulldozers*. In the forward direction, the vehicle term *bulldozer* suggests a category of forceful behavior that can provide values for the dimensions of persuasiveness, need for sales, and so forth that characterize salesmen. In Glucksberg et al.'s account, the superiority of this forward direction over the reversed direction (i.e., *Some bulldozers are salesmen*) arises from role-specific processes that hold throughout processing. Either *salesmen* fails to provide a possible category, *bulldozer* fails to provide possible dimensions, or the two do not link in a manner sufficiently salient to initiate early processing. On this account, role-specific metaphoric processing cannot be initiated in the reversed direction.

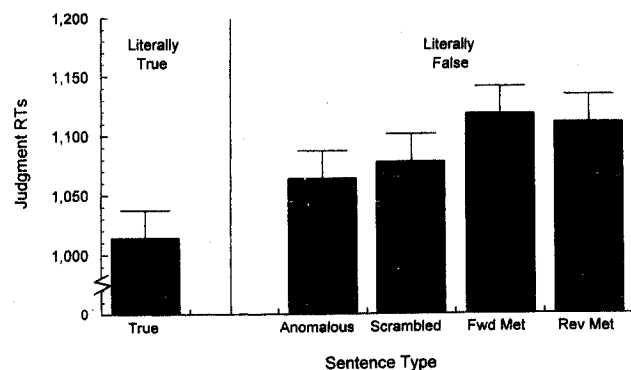


Figure 1. Mean judgment reaction times (RTs) for literally true and anomalous statements and for scrambled, forward (Fwd Met), and reversed (Rev Met) metaphors with associated 95% confidence intervals.

Therefore, it should have been easy to reject reversed metaphors and to correctly classify them as literally false.

The finding that reversed metaphors create as much interference as do forward metaphors poses problems for the property attribution account and supports the claim that metaphor processing begins with an initial alignment process. However, one might question the generality of these findings. Several researchers have noted processing differences between high-conventional metaphors, such as *A rumor is a virus*, and low-conventional metaphors, such as *A detective is a ferret* (Blasko & Connine, 1993; Bowdle & Gentner, 1995, 1999a, 1999b; Gentner & Wolff, 1997, 2000; Giora, 1997; Wolff & Gentner, 1992). The property attribution process may apply best to high-conventional metaphors, that is, to metaphors whose vehicle terms have associated with them—typically as a secondary word sense—the abstraction intended by the metaphor (e.g., “contagious and harmful” for the metaphor *A rumor is a virus*). For example, Gentner and Wolff (1997) found that comprehension time for conventional metaphors was lower if participants were primed by seeing the vehicle term immediately before the metaphor. This effect was not found for novel metaphors. This suggests that early access to the abstraction associated with the vehicle or base term is helpful in processing conventional metaphors. This would follow if metaphoric bases gradually develop alternative senses as they become conventionalized (Bowdle & Gentner, 1995, 1999b; Gentner & Wolff, 1997), resulting finally in the development of a dual representation of the vehicle or base term. The implication is that directional (role specific) processing may be most pronounced for high-conventional metaphors. Therefore, to ensure a fair test of the directional asymmetry position, we decided to use only highly conventional metaphors in the next study.

Another factor that may influence the results is the *metaphorical similarity* between the base and the target (vehicle and topic) terms. By metaphorical similarity we do not mean overall similarity, which tends to undermine metaphorical aptness (Ortony, 1979; Tourangeau & Sternberg, 1981). Rather, metaphorical similarity refers to the degree to which the two concepts share the schemas or properties conveyed by the metaphorical assertion. According to structural alignment theory, when the base and the target share much of the common system that the metaphor conveys, alignment should be easier than for a low-similarity metaphor. Thus, high-similarity metaphors such as *That soldier is a pawn* should be faster to process than low-similarity metaphors such as *That senator is a pawn*. This is because the abstraction associated with *pawn*—“being controlled from above”—is also present to some degree in representations of *soldier*, resulting in a rather large match. In contrast, this abstraction is not prominent in the representation of *senator*, resulting in a rather sparse alignment (and in a subsequent large-scale projection from *pawn* to *senator*). Consistent with this prediction, Gentner and Wolff (1997) found that high-similarity metaphors were faster to process than were low-similarity metaphors.

But if high-similarity metaphors are especially conducive to rapid alignment processing, as Gentner and Wolff (1997) predicted, then such metaphors may not be a fair test of asymmetrical processing in metaphor. There could be more than one way of processing metaphor. Perhaps high-similarity metaphors are processed through alignment, but other metaphors are processed through property attribution, so that here the initial process is role specific. In this case, we would expect to see greater initial asymmetry for low-similarity metaphors. To test this possibility, and to give initially asymmetrical models a fair trial, we manipulated similarity; Experiment 2 included both low-similarity and high-similarity metaphors.

Experiment 2 used the same logic as Experiment 1. We again used the metaphor interference technique with forward and reversed metaphors to test whether early metaphor processing is symmetrical (role neutral) or asymmetrical (role specific). The new features of this study were that (a) the metaphors were all of high conventionality and (b) the metaphorical similarity of the vehicle and the topic was manipulated to be either high or low. The predictions are as follows: The attributive categorization model predicts (a) interference only for forward metaphors and (b) either interference for all metaphors or more interference for low-similarity metaphors. The structure-mapping model predicts (a) role-independent interference and (b) more interference for high-similarity metaphors than for low-similarity metaphors.

## Experiment 2

### Method

**Participants.** The participants were 32 Northwestern undergraduates who participated for course credit.

**Materials.** The materials were similar to those of Experiment 1 except that the metaphors were selected to be of high conventionality and to vary in vehicle–topic similarity.<sup>1</sup> As before, the metaphors were highly directional, as indicated by a directionality task discussed in this section. Four 96-item test lists were constructed, each containing 48 literally true statements (24 high typical, 24 low typical), 16 anomalous (literally false) statements, 16 metaphors (4 forward high similarity, 4 forward low similarity, 4 reversed high similarity, and 4 reversed low similarity), and 16 scrambled metaphors. The metaphors used are shown in Appendix B. On each list, half of the metaphors were forward and half were reversed. For both directions, half had high-similarity topics and the other half had low-similarity topics. We accomplished this by having two topic terms for each base, one of high similarity and the other of low similarity. Thus, there were six kinds of literally false statements: (a) anomalous statements, (b) forward high-similarity metaphors (e.g., *Some desks are junkyards*), (c) forward low-similarity metaphors (e.g., *Some tables are junkyards*), (d) reversed high-similarity metaphors (e.g., *Some junkyards are desks*), (e) reversed low-similarity metaphors (e.g., *Some junkyards are tables*), and (f) scrambled metaphors. As in Experiment 1, we constructed anomalous statements by recombining the subject and predicate

<sup>1</sup> The metaphors used here were taken from a larger set rated by Gentner and Wolff (1997). A more complete description of the rating tasks is given in that article.

nouns of the literally true statements and constructed scrambled metaphors by recombining the subject and predicate nouns of the metaphors. A 60-item practice list was constructed with 30 literally true (15 high typical, 15 low typical) and 30 anomalous statements. Both the practice and the test lists were individually randomized for each participant. Preceding the test lists were 12 warm-up (filler) items, 6 true and 6 anomalous.

**Directionality preference task.** The method for obtaining directionality preferences was the same as in Experiment 1. The forward direction was chosen, on average, 92% of the time, with responses to individual items varying from 69% to 100%.

**Metaphorical similarity and conventionality.** Twenty-four Northwestern undergraduates, using a 1 to 7 scale, rated metaphorical or relational similarity, explained as follows: "Things are relationally similar when they participate in the same relationships. For example, a cigarette and a time bomb are relationally similar because they both can cause harm after a period of apparent harmlessness. . . ." The high-similarity metaphors had a mean rating of 3.57, and the low-similarity metaphors, 2.34. Conventionality ratings were collected as follows. First, 10 Northwestern undergraduates wrote interpretations for a set of 64 metaphors. For each metaphor, the interpretation that was most frequently provided was rated for vehicle conventionality on a 1 to 7 scale by 20 Northwestern undergraduates, using the following format: "When we say something is a rocket, how conventional is the interpretation that this is something that moves very fast?" The mean conventionality of the metaphors used in the present study was 6.02.

**Procedure.** The procedure was the same as in Experiment 1.

**Design.** A within-subject design of similarity (high or low) and direction (forward or reversed) was used, with each participant verifying each sentence type. The assignment of metaphor direction and similarity level to particular sets of metaphors was counterbalanced across four between-subjects groups.

## Results and Discussion

Incorrect responses (9%) and RTs greater than 2,400 ms (3%) were removed, resulting in the exclusion of 12% of the data. The results, shown in Table 1, fit the predictions of structure-mapping theory. When metaphorical similarity was high, interference effects were found for metaphors in both forward and reversed directions, replicating the results in Experiment 1. Also consistent with structure-mapping

was the fact that interference effects were greater for high-similarity metaphors than for low-similarity metaphors; indeed, no interference effects were found for low-similarity metaphors in either forward or reversed directions. Contrary to the predictions of the attributive categorization model, no evidence for role-specific interference was found, even for high-conventional metaphors, regardless of similarity.

These observations are supported by statistical analyses focusing on the six different kinds of false statements. A one-way repeated measures ANOVA indicated an overall effect of sentence type,  $F_s(5, 155) = 2.56, p < .05, MSE = 20,092$ ;  $F_i(5, 90) = 2.47, p < .05, MSE = 19,532$ . To further examine the results, we conducted individual comparisons, based on a pooled error term, assuming a .01 significance level derived from the Bonferroni method.

As in Experiment 1, forward and reversed metaphors were strikingly similar in their behavior; regardless of direction, high-similarity metaphors led to interference, and low-similarity metaphors did not.

High-similarity metaphors showed the same pattern as was found in Experiment 1. More important, both forward ( $M = 1,266$  ms) and reversed ( $M = 1,254$  ms) high-similarity metaphors took longer to reject than did scrambled metaphors ( $M = 1,175$  ms),  $t_s(31) = 3.63, p < .001$ ;  $t_i(30) = 3.23, p < .01$ ; and  $t_s(31) = 3.15, p < .01$ ;  $t_i(30) = 2.69, p = .012$ . Also as in Experiment 1, scrambled versions of these metaphors did not differ from anomalous statements ( $M = 1,183$  ms),  $t_s(31) = 0.48, ns$ ;  $t_i(30) = 0.54, ns$ . The fact that scrambled metaphors did not give rise to interference again reinforces the claim that the interference effects found for high-similarity forward and reversed metaphors reflect interactions between the base and the target, and not effects specific to the particular words.

A key prediction of the structure-mapping model is that high-similarity metaphors should lead to more interference than low-similarity metaphors, regardless of direction. To test this prediction, we analyzed metaphors in a  $2 \times 2$  ANOVA with the factors being similarity (high, low) and direction (forward, reversed). As expected, high-similarity metaphors took longer to reject than did low-similarity metaphors, significantly so across subjects,  $F_s(1, 31) = 6.09, p < .05$ , and marginally so across items,  $F_i(1, 60) = 3.03, p = .087$ . Although low-similarity, high-conventional metaphors seemed the ideal candidates to display directed processing (i.e., differences between forward and reversed orderings), such effects were not found. There was no overall effect of direction,  $F_s(1, 31) = 0.012, ns$ ;  $F_i(1, 60) = 0.008, ns$ ; nor did direction interact with similarity,  $F_s(1, 31) = 0.105, ns$ ;  $F_i(1, 60) = 0.153, ns$ . The low-similarity metaphors—both forward and reversed—did not differ from scrambled metaphors,  $t_s(31) = 0.40, ns$ , and  $t_i(31) = 0.64, ns$ ;  $t_i(30) = 0.80, ns$ , and  $t_i(30) = 1.17, ns$ . In combination, the results suggest that low-similarity metaphors failed to produce interference in either direction and that high-similarity metaphors produced interference in both directions.

Table 1  
Mean Reaction Time (RT) in Milliseconds as a Function of Sentence Type in Experiment 2

Sentence type, with examples	n	M	SD	Error rate (%)
Literally true	48	1,141	197	11
Anomalous	16	1,183	217	13
Scrambled ( $n = 16$ )	16	1,175	226	1
High-similarity forward metaphors				
Some suburbs are parasites.	16 across P	1,266	271	12
High-similarity reversed metaphors				
Some parasites are suburbs.	16 across P	1,254	302	6
Low-similarity forward metaphors				
Some towns are parasites.	16 across P	1,185	264	8
Low-similarity reversed metaphors				
Some parasites are towns.	16 across P	1,191	258	5

Note. P = Participants.

Overall, the evidence provides support for early role-neutral processing of metaphors. However, these conclusions are based on the assumption that the metaphor interference technique does, in fact, tap early processing. It is conceivable that the results instead reflect late processes. For example, perhaps participants quickly recognized that the reversed metaphors were uninterpretable and reversed them. There are some considerations that argue for the early-processing assumption. First, it's not clear why participants would seek to mentally transform the reversed metaphors, given that their task is one of true-false literal judgment. Second, the metaphor interference technique does not require that processing be completed; participants should reject the metaphors as soon as they sense that a literal interpretation is not possible. Thus, the interference seems likely to be associated with the initial stages of metaphor processing.

Perhaps the most important reason to believe that these results reflect early rather than late processing is that studies of metaphor comprehension typically show much higher processing times than those associated with the metaphor interference phenomenon. In metaphor comprehension studies, the mean RTs typically lie between 1,800 and 4,000 ms (e.g., Gentner & Wolff, 1997: 3,798 ms; Glucksberg et al., 1997: 2,292 ms; Gregory & Mergler, 1990: 3,480 ms). In contrast, metaphor interference studies show mean RTs ranging between 1,000 and 1,400 ms (e.g., Glucksberg et al., 1982: 1,369 ms; Keysar, 1989: 1,025 ms; and in our current studies, 1,115 ms in Experiment 1 and 1,224 ms in Experiment 2). These times are considerably lower than those found in full comprehension studies and are consistent with the view that the metaphoric interference effect taps into an early stage of processing.

However, to be sure of our interpretation, we decided to test the late-stage reversal possibility directly by giving another group of participants the same set of metaphors in a standard metaphor comprehension task.<sup>2</sup> This also allowed for a more complete test of our predictions. In Experiment 3, participants received the same set of metaphors that were used in Experiment 2 and then judged whether the metaphor was comprehensible. There were three main predictions. First, the comprehensibility judgment should require more time than did the literal true-false judgment in Experiment 2 for the same metaphors. This is because the comprehensibility task invites participants to carry through the whole task of metaphor processing. This finding is of course essential to our interpretation of the metaphor interference results as indicative of early processing. Second, participants should be sensitive to role information in their comprehensibility judgments; forward metaphors should be rated as more comprehensible than reversed metaphors. This pattern is important in confirming that the metaphors were indeed directional but that the role specificity did not emerge in early processing. Third, structure-mapping theory predicts that high-similarity metaphors should be more readily aligned, and thus more likely to be judged comprehensible, than low-similarity metaphors.

Table 2

*Comprehensibility Ratings as a Function of Sentence Type in Experiment 3*

Sentence type	n	Comprehensible (%)
Scrambled	16	6
High-similarity forward metaphors	16 across P	81
High-similarity reversed metaphors	16 across P	46
Low-similarity forward metaphors	16 across P	69
Low-similarity reversed metaphors	16 across P	29

Note. P = Participants.

### Experiment 3

#### Method

**Participants.** The participants were 40 Northwestern undergraduates who participated for course credit.

**Materials.** We constructed four 32-item test lists using the same materials employed in Experiment 2, minus the literal class-inclusions (see Appendix B).<sup>3</sup> Each list contained 16 metaphors and 16 scrambled metaphors. Half of the metaphors were forward and half were reversed. As in Experiment 2, in both directions, half of the metaphors had high-similarity topics and the other half had low-similarity topics. A 32-item practice list was constructed with 16 metaphors and 16 anomalous statements.

**Procedure.** Participants were told that they would see either metaphorical statements, like *Some antiques are fossils*, or anomalous statements, like *Some balloons are cactuses*. Their task was to say whether these statements were comprehensible. Participants were then given 20 practice trials pressing the left and right arrow keys in response to the phrases "Comprehensible" or "Incomprehensible," respectively. The participants then received 64 practice trials with sample metaphoric statements and finally 72 test trials. Presentation of the sentences in both the practice and the test sessions was as in previous experiments except that participants were informed of incorrect responses only during the practice session, not during the test session, as was the case in previous experiments. This was done in order to allow participants to make their own decisions as to the comprehensibility of the reversed metaphors. Participants were notified by a message and a soft bell that their responses were "Too slow" after 4,000 ms.

**Design.** A within-subject design of Similarity  $\times$  Direction was used, with each participant judging each sentence type within a single session. The assignment of metaphor direction (forward vs. reversed) and similarity level (high vs. low) was counterbalanced across four between-subjects groups. Order of the items was randomized across participants.

#### Results and Discussion

RTs greater than 3,400 ms were removed, resulting in the exclusion of 3% of the data. The results, shown in Table 2, were as predicted. First, comparing the results with those from Experiment 2 (metaphors only), we found that comprehensibility decisions ( $M = 1,687$  ms,  $SD = 360$ ) took longer

<sup>2</sup> We thank Sam Glucksberg for raising this issue.

<sup>3</sup> Literal class-inclusion statements were not used in this experiment to forestall participants' reinterpreting of the comprehensibility measure as a decision about literal truth and falsity.



than literal truth or falsity decisions ( $M = 1,224$  ms,  $SD = 237$ ). A test between these two means was statistically reliable both across participants,  $t_s(55) = 5.84$ ,  $p < .001$ , and across items,  $t_i(23) = 7.46$ ,  $p < .001$ . Of course, such a result must be interpreted cautiously because it compares data from different experiments.

The chief predictions concerned the comprehensibility decisions. According to the structure-mapping model, an initial alignment process is followed by directional projection of inferences. Thus metaphoric directionality should emerge if people are allowed to complete the comprehension process. This is exactly what was found. Participants were roughly twice as likely to judge forward metaphors ( $M = .75$ ,  $SD = .19$ ) to be comprehensible as reversed metaphors ( $M = .37$ ,  $SD = .275$ ),  $F_s(1, 39) = 28.9$ ,  $p < .001$ ;  $F_i(1, 15) = 132.6$ ,  $p < .001$ . In conjunction with the finding that the comprehensibility RTs are longer than the interference RTs, these findings suggest that directional processes occur in metaphor comprehension but that they emerge relatively late.

The third prediction of the structure-mapping model is that high-similarity metaphors should be more comprehensible than low-similarity metaphors. Indeed, participants considered metaphors that were high in similarity ( $M = .63$ ,  $SD = .219$ ) more comprehensible than metaphors that were low in similarity ( $M = .49$ ,  $SD = .197$ ),  $F_s(1, 39) = 28.9$ ,  $p < .001$ ;  $F_i(1, 15) = 6.67$ ,  $p < .05$ . Also consistent with this prediction was the fact that participants were more likely to judge both forward and reversed metaphors ( $M = .56$ ,  $SD = .190$ ) to be comprehensible than scrambled metaphors ( $M = .06$ ,  $SD = .096$ ),  $t_s(39) = 14.90$ ,  $p < .001$ ;  $t_i(30) = 10.61$ ,  $p < .001$ . All other factors and interactions were nonsignificant.

In sum, the results confirm that full metaphor processing is role sensitive, as most theories of metaphor would predict. They also support the assumption that metaphor interference effects tap early processing of metaphors. When people are asked to comprehend metaphors, they spend more time than they do when they make a literal true-false judgment. These results support the claim that role sensitive processes occur late in processing. Metaphor interference effects occur early and are role neutral, whereas full comprehension requires more time and is directional.

### General Discussion

The results of this research support the claim that early metaphoric processing is symmetrical and that later processing is directional, as predicted by structure-mapping theory. In Experiment 1, we replicated the metaphor interference effects of Glucksberg et al. (1982) and Keysar (1989), indicating rapid initial processing of metaphors. Although the metaphors were chosen to be clearly directional, these interference effects were independent of the order of the vehicle and topic terms. These findings suggest that metaphoric processing is initially role neutral (symmetrical). In Experiment 2, we again tested for early asymmetry, this time using metaphors that seemed particularly apt for revealing role-specific processing—namely, high-conventional meta-

phors with low base-target (vehicle-topic) metaphorical similarity—along with high-similarity metaphors for comparison. Again, we found no difference between forward and reversed metaphors. When metaphorical similarity was high, order-independent interference was found, as in the first experiment. When metaphorical similarity was low, there were no interference effects for either forward or reversed metaphors. In Experiment 3, we verified that the metaphors were directional when processed to completion. When asked to make comprehensibility judgments, people were about twice as likely to judge forward metaphors to be comprehensible as they were to judge reversed metaphors. Experiment 3 also verified that full comprehension requires considerably longer times than those associated with metaphor interference in Experiments 1 and 2, consistent with the assumption that metaphor interference results from an early stage in metaphor processing.

### Directionality in Metaphor Processing

Metaphors are clearly directional; there is abundant evidence that reversing metaphors affects their perceived aptness and interpretability (Gentner & Clement, 1988; Glucksberg et al., 1997; Ortony, 1979; Ortony et al., 1985). But when and how does this asymmetry come about? There are at least three possibilities. The strongest is that there is an initial *temporal processing* asymmetry: Processing begins with the abstraction and projection of a category from the vehicle. This model seems consistent with the class-inclusion account put forward by Glucksberg and Keysar (1990; though it should be noted that they did not specify a particular processing model). In addition to the present results, there is evidence against such a temporal asymmetry model from priming studies by Gentner and Wolff (1997, 2000) and by Wolff and Gentner (1992). If processing begins with category projection from the vehicle, metaphors should be comprehended more readily when primed by the vehicle or base term than when primed by the topic or target term. However, these studies revealed no such base advantage. In general, metaphors were comprehended no faster when preceded by the base than when preceded by the target. These findings, as well as the present findings, run strongly against the temporal asymmetry account.

A second account of metaphoric directionality, Glucksberg et al.'s (1997) attributive categorization model, states that there is a *processing* asymmetry. Processing begins simultaneously with both terms but is differentiated from the start in role-specific ways. The vehicle term provides the attributive category that will be used to characterize the topic, and the topic provides the attributional dimensions along which it can be characterized. As evidence for role-specific processing, Glucksberg et al. showed that metaphors were faster to be comprehended when preceded by unambiguous vehicles—vehicles that uniquely exemplify a particular attributive category—than by ambiguous vehicles, and by high-constraint topics—topics that produce limited expectations about the attributive dimensions along which they might be characterized—than by low-constraint topics.



A concern here is that degree of ambiguity and degree of constraint could simply be interrelated aspects of the informativity of their terms, rather than being separate, role-specific dimensions. Glucksberg et al. (1997) countered this possibility by carrying out a rating study on their materials that indicated that the two dimensions were orthogonal. However, establishing the claim of role-specific processing also requires a demonstration that the two kinds of facilitation are specific to their particular roles. That is, it must be shown that (a) high constraint is facilitative for topics but (b) it is not facilitative (or significantly less so) for vehicles, and that (c) low ambiguity is facilitative for vehicles but (d) is not facilitative (or significantly less so) for topics. In their studies, Glucksberg et al. tested only the positive claims (a and c) and not the negative claims (b and d). Thus it remains possible that the advantage derived from the general informativity of the primes, rather than from specific processing affordances.

The present results argue against role-specific processing in the initial stages (see Wolff & Gentner, 1999, for further evidence). This seems to leave us with the third possibility, that metaphoric asymmetry emerges as the result of a directional process of inference projection that follows an initial role-neutral stage, as predicted by the structure-mapping account.

### *Linking Figurative With Literal Processing*

The traditional separation between literal and figurative processing has been questioned of late. First, the view that metaphors are processed in two stages—a literal stage followed by additional processing when literal interpretation fails (Clark & Lucy, 1975; Searle, 1979)—has failed to garner consistent empirical support. As reviewed by Hoffman and Kemper (1986) and by Gibbs and Gerrig (1989), metaphors do not appear to require more time than do literal comparisons, as would be expected if they involved more processing stages (Inhoff, Lima, & Carroll, 1984; Ortony, Schallert, Reynolds, & Antos, 1978; however, see Janus & Bever, 1985). Further, as replicated here, metaphoric processing does not appear to be optional: Participants told to focus only on literal meaning were nonetheless unable to ignore metaphoric interpretations (Gildea & Glucksberg, 1983; Glucksberg et al., 1982; Keysar, 1989). These findings have led some researchers to suggest that literal and metaphoric interpretations may involve the same processes (Bowdle & Gentner, 1997; Falkenhainer et al., 1989; Gentner, 1989; Gentner & Bowdle, 1994; Gentner & Clement, 1988; Gentner, Falkenhainer, & Skorstad, 1988; Gibbs, 1994; Gibbs & Gerrig, 1989; Medin, Goldstone, & Gentner, 1993; Ortony, 1979; Ortony et al., 1978; Rumelhart, 1979; Verbrugge, 1977).

Structure-mapping theory offers a way in which metaphorical and literal comparisons can be captured within a single mechanism. It is perfectly possible to arrive at both a literal and a metaphorical interpretation for the same comparison; the process model is indifferent to this distinction. For example, on hearing *My neighbor is a butcher*, one might simultaneously entertain the possibility that he cuts meat for

a living and the possibility that he trims hedges in a clumsy and brutal manner. The interpretation that would win out would be determined by aptness, informativity, and contextual fit. Murphy (1996) noted that for many metaphors, both the vehicle and the topic could potentially participate in several different metaphorical meanings. Structural alignment can act as a constraint on this selection problem.

A study by Onishi and Murphy (1993) on metaphoric reference produced further evidence consistent with the structural alignment position. According to the alignment view, metaphor processing should be facilitated when the two terms being compared are close in the text. Onishi and Murphy found that metaphors were understood as quickly as literals when the two terms were located in the same sentence (e.g., *That boxer is a creampuff*); however, metaphors were harder to interpret than were literal comparisons when the two terms were expressed in separate sentences: specifically, when the vehicle was used as an anaphoric reference for the topic (e.g., *The creampuff didn't even show up*).

### *A Unified Process Model*

One appealing feature of the structure-mapping approach is its assumption of a common set of processes for literal and nonliteral comparisons, and for both metaphor and analogy. This provides a framework for discussing such phenomena as extended metaphors (Gentner & Boronat, 1992; Gentner & Wolff, 2000; Lakoff & Johnson, 1980; McGlone & Bortfeld, in press), such as mappings from space to time (Gentner, in press; Gentner & Imai, 1992; McGlone & Harding, 1998). There is empirical support for structural alignment in the processing of similarity, as well as the processing of analogy and metaphor (Bowdle & Gentner, 1997; Gentner & Markman, 1993, 1994, 1995, 1997; Gentner & Rattermann, 1991; Goldstone, 1994; Goldstone & Medin, 1994; Goldstone, Medin, & Gentner, 1991; Markman & Gentner, 1993a, 1993b, 1996; Medin et al., 1993).

### *Implications for Models of Analogy*

Models of analogical processing, like models of metaphor, differ in whether they begin with a directed projection process followed by matching and verification or with a symmetric matching process followed by directed projection (Gentner, Bowdle, Wolff, & Boronat, in press). The former class includes explanation-based-learning models of analogy (Kedar-Cabelli, 1988), abstraction-based models (Burstein, 1983; Carbonell, 1982, 1986; Greiner, 1988), models in which statements are incrementally mapped from the base to the target (Keane & Brayshaw, 1988; Keane, Ledgeway, & Duff, 1994), and models such as Hummel and Holyoak's (1997) Learning and Inference with Schemas and Analogies (LISA) model,<sup>4</sup> in which information is projected from a

<sup>4</sup> However, this model allows for shifting between the analogues as to which is driver and which is recipient, so its predictions are not entirely clear.

driver to a recipient. The latter class includes alignment-first models (Falkenhainer et al., 1989; Holyoak & Thagard, 1989). The current findings of initial role-neutral processing provide support for alignment-based models of analogy as well as metaphor.

### *The Career of Metaphor*

We have recently proposed an account of the conventionalization of metaphor that we call the *career of metaphor* hypothesis: Novel metaphors are processed as structural alignments between the concrete or literal representations of the base and target, but as repeated comparisons are made, the metaphorical meaning is gradually abstracted and comes to be associated with the vehicle or base term (Bowdle & Gentner, 1995, 1999a, 1999b; Gentner & Wolff, 1997, 2000). The metaphoric meaning may eventually become the primary sense, as in *phlegmatic* or *sanguine*. This account is consistent with research suggesting that conventional metaphors (Blank, 1988; Blasko & Connine, 1993; Giora, 1997; Lakoff, 1990; Lehrer, 1990) and idioms (Cacciari & Tabossi, 1988; Gibbs, 1980, 1994; Swinney & Cutler, 1979) have stock lexicalized meanings associated with their bases.

This account is supported by the results of priming studies. As discussed previously, our studies failed to show any general advantage for base priming over target priming in the time required to process a metaphor (Gentner & Wolff, 1997; Wolff & Gentner, 1992). However, there was one exception: A base advantage was found for metaphors with highly conventional bases and low metaphoric or relational similarity between base and target (e.g., *Some towns are parasites*). Gentner and Wolff proposed that metaphoric processing may involve both alignment and category inheritance. Alignment is fostered by base-target similarity, and category inheritance is fostered by base conventionality. This last process occurs because conventional bases have developed secondary metaphoric meanings, and here our account converges with Glucksberg and Keysar's (1990) proposal that metaphors have dual representations. (The difference is that Glucksberg and Keysar proposed that dual representation is a universal feature of metaphors, whereas we propose that it is characteristic only of conventional metaphors. Further, they proposed dual representation as an alternative to comparison and alignment, whereas we propose that the metaphoric representations arise out of repeated alignments.) The present findings suggest that alignment processing dominates for high-similarity metaphors, whether novel or conventional. However, for metaphors with highly conventional bases and neutral targets (i.e., low-similarity targets), category projection is the dominant process.

### *Summary*

Our results indicate initial role-neutral processing of metaphors. We found no evidence for early role sensitivity in the processing of metaphors, regardless of similarity or conventionality. The evidence here is most consistent with the predictions of structure-mapping theory: Metaphoric

processing begins as an alignment between the target and base representations, followed by directional projection of inferences.

### References

- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. *Journal of Experimental Psychology Monographs*, 80, 2-45.
- Blank, G. D. (1988). Metaphors in the lexicon. *Metaphor and symbolic activity*, 3, 21-36.
- Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 295-308.
- Bowdle, B., & Gentner, D. (1995, November). *The career of similarity*. Paper presented at the meeting of the Psychonomics Society, Los Angeles, CA.
- Bowdle, B., & Gentner, D. (1997). Informativity and asymmetry in comparisons. *Cognitive Psychology*, 34, 244-286.
- Bowdle, B., & Gentner, D. (1999a, August). *Metaphor comprehension: From comparison to categorization*. Paper presented at the 21st Annual Meeting of the Cognitive Science Society, Vancouver, British Columbia.
- Bowdle, B., & Gentner, D. (1999b). *The career of metaphor*. Manuscript in preparation.
- Brysbaert, M., Bovens, N., & d'Ydewalle, G. (1989). Turbo Pascal timing routines for the IBM microcomputer family. *Behavior Research Methods, Instruments, & Computers*, 21, 73-83.
- Burstein, M. H. (1983). Concept formation by incremental analogical reasoning and debugging. In R. Michalski, J. Carbonell, & T. Mitchell (Eds.), *Machine learning* (Vol. 2, pp. 351-369). Los Altos, CA: Kaufmann.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory & Language*, 27, 668-683.
- Carbonell, J. G. (1982). Metaphor: An inescapable phenomenon in natural language comprehension. In W. G. Lehnert & M. H. Ringle (Eds.), *Strategies for natural language processing* (pp. 415-435). Hillsdale, NJ: Erlbaum.
- Carbonell, J. G. (1986). Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2, pp. 371-392). Los Altos, CA: Kaufmann.
- Clark, H. H., & Lucy, P. (1975). Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 12, 335-359.
- Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science*, 15, 89-132.
- Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: An algorithm and examples. *Artificial Intelligence*, 41, 1-63.
- Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19, 141-205.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). New York: Cambridge University Press.
- Gentner, D. (in press). Spatial metaphors in temporal reasoning. In M. Gattis (Ed.), *Spatial schemas in abstract thought*. Cambridge, MA: MIT Press.
- Gentner, D., & Boronat, C. B. (1992). *Metaphor as mapping*. Paper presented at the Workshop on Metaphor, Tel Aviv, Israel.

- Gentner, D., & Bowdle, B. F. (1994). The coherence imbalance hypothesis: A functional approach to asymmetry in comparison. In A. Ram & K. Elselst (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 351–356). Hillsdale, NJ: Erlbaum.
- Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. B. (in press). Metaphor is like analogy. In D. Gentner, K. J. Holyoak, & B. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.
- Gentner, D., & Clement, C. A. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. In G. H. Bower (Ed.), *The psychology of learning and motivation, advances in research and theory* (Vol. 22, pp. 307–358). New York: Academic Press.
- Gentner, D., Falkenhainer, B., & Skorstad, J. (1988). Viewing metaphor as analogy. In D. H. Herman (Ed.), *Analogical reasoning* (pp. 171–177). New York: Kluwer Academic.
- Gentner, D., & Imai, M. (1992). Is the future always ahead? Evidence for system-mappings in understanding space-time metaphors. *Proceedings of the Fourteenth Annual Meeting of the Cognitive Science Society* (510–515). Hillsdale, NJ: Erlbaum.
- Gentner, D., & Markman, A. B. (1993). Analogy—Watershed or Waterloo? Structural alignment and the development of connectionist models of cognition. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems* (Vol. 5, pp. 855–862). San Mateo, CA: Kaufmann.
- Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science*, 5, 152–158.
- Gentner, D., & Markman, A. B. (1995). Similarity is like analogy. In C. Cacciari (Ed.), *Similarity in language, thought, and perception* (pp. 111–148). Brussels, Belgium: BREPOLs.
- Gentner, D., & Markman, A. B. (1997). Structure-mapping in analogy and similarity. *American Psychologist*, 52, 45–56.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Brynes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225–277). London: Cambridge University Press.
- Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory & Language*, 37, 331–355.
- Gentner, D., & Wolff, P. (2000). Metaphor and knowledge change. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual change in humans and machines* (pp. 295–342). Mahwah, NJ: Erlbaum.
- Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8, 449–456.
- Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge, England: Cambridge University Press.
- Gibbs, R. W., & Gerrig, R. J. (1989). How context makes metaphor comprehension seem “special.” *Metaphor and Symbolic Activity*, 4, 145–158.
- Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior*, 22, 577–590.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8, 183–206.
- Glucksberg, S., Gildea, P., & Bookin, H. (1982). On understanding non-literal speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, 1, 85–96.
- Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, 97, 3–18.
- Glucksberg, S., & Keysar, B. (1993). How metaphors work. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 447–480). Cambridge, England: Cambridge University Press.
- Glucksberg, S., McGlone, M. S., & Manfredi, D. (1997). Property attribution in metaphor comprehension. *Journal of Memory & Language*, 36, 50–67.
- Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 3–28.
- Goldstone, R. L., & Medin, D. L. (1994). Time course of comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 29–50.
- Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology*, 23, 222–264.
- Gregory, M. E., & Mergler, N. L. (1990). Metaphor comprehension: In search of literal truth, possible sense, and metaphoricity. *Metaphor and Symbolic Activity*, 5, 151–173.
- Greiner, R. (1988). Abstraction-based analogical inference. In D. H. Herman (Ed.), *Analogical reasoning* (pp. 147–170). New York: Kluwer Academic.
- Hoffman, R. R., & Kemper, S. (1986). What could reaction-time studies be telling us about metaphor comprehension? *Metaphor and Symbolic Activity*, 2, 149–186.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295–355.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, 104, 427–466.
- Inhoff, A. W., Lima, S. D., & Carroll, P. J. (1984). Contextual effects on metaphor comprehension in reading. *Memory & Cognition*, 12, 558–567.
- Janus, R. A., & Bever, T. G. (1985). Processing of metaphoric language: An investigation of the three-stage model of metaphor comprehension. *Journal of Psycholinguistic Research*, 14, 473–487.
- Keane, M. T., & Brayshaw, M. (1988). The incremental analogical machine: A computational model of analogy. In D. Sleeman (Ed.), *Third European working session on machine learning* (pp. 53–62). San Mateo, CA: Kaufmann.
- Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science*, 18, 387–438.
- Kedar-Cabelli, S. (1988). Analogy—From a unified perspective. In D. H. Helman (Ed.), *Analogical reasoning: Perspectives of artificial intelligence, cognitive science, and philosophy* (pp. 65–103). Dordrecht, the Netherlands: Kluwer Academic.
- Keysar, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory & Language*, 28, 275–385.
- Lakoff, G. (1990). The invariance hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics*, 1, 39–74.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lehrer, A. (1990). Polysemy, conventionality, and the structure of the lexicon. *Cognitive Linguistics*, 1, 207–246.
- Markman, A. B., & Gentner, D. (1993a). Splitting the differences: A structural alignment view of similarity. *Journal of Memory & Language*, 32, 517–535.
- Markman, A. B., & Gentner, D. (1993b). Structural alignment during similarity comparisons. *Cognitive Psychology*, 23, 431–467.
- Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition*, 24, 235–249.

- McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, 11, 1-37.
- McGlone, M. S., & Bortfeld, H. (in press). The continuum of metaphor processing. *Proceedings of the AISB '99 Conference on Metaphor, Artificial Intelligence, and Cognition*.
- McGlone, M. S., & Harding, J. L. (1998). Back (or forward?) to the future: The role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1211-1223.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100, 254-278.
- Miller, G. A. (1993). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 357-400). Cambridge, England: Cambridge University Press.
- Murphy, G. L. (1996). On metaphoric representation. *Cognition*, 60, 173-204.
- Onishi, K. H., & Murphy, G. L. (1993). Metaphoric reference: When metaphors are not understood as easily as literal expressions. *Memory & Cognition*, 21, 763-772.
- Ortony, A. (1979). Beyond literal similarity. *Psychological Review*, 86, 161-180.
- Ortony, A., Schallert, D., Reynolds, R., & Antos, S. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, 17, 467-477.
- Ortony, A., Vondruska, R. J., Foss, M. A., & Jones, L. E. (1985). Salience, similes, and the asymmetry of similarity. *Journal of Memory & Language*, 24, 569-594.
- Rumelhart, D. E. (1979). Some problems with the notion of literal meanings. In A. Ortony (Ed.), *Metaphor and thought* (pp. 78-123). Cambridge, England: Cambridge University Press.
- Searle, J. (1979). Metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 92-123). Cambridge, England: Cambridge University Press.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81, 214-241.
- Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18, 523-534.
- Tourangeau, R., & Sternberg, R. (1981). Aptness in metaphor. *Cognitive Psychology*, 13, 27-55.
- Verbrugge, R. R. (1977). Resemblances in language and perception. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing* (pp. 365-392). Hillsdale, NJ: Erlbaum.
- Wolff, P., & Gentner, D. (1992). The time course of metaphor comprehension. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 504-509). Hillsdale, NJ: Erlbaum.
- Wolff, P., & Gentner, D. (1999). A two-stage model of metaphor processing. Manuscript submitted for publication.

## Appendix A

### Metaphors Used in Experiment 1

---

Some jobs are jails.  
 Some vacations are doctors.  
 Some surgery is a balancing act.  
 Some laughter is therapy.  
 Some groupies are satellites.  
 Some slums are tumors.  
 Some auditions are doors.  
 Some lectures are lullabies.  
 Some contracts are handcuffs.  
 Some diplomas are passports.  
 Some rumors are viruses.  
 Some salesman are bulldozers.  
 Some gossips are radios.  
 Some suburbs are parasites.  
 Some students are sponges.  
 Some ships are plows.  
 Some greed is a demon.  
 Some books are compasses.  
 Some consciences are harnesses.  
 Some promises are gambles.  
 Some debts are weights.  
 Some yawning is a plague.  
 Some fame is a beacon.  
 Some libraries are goldmines.  
 Some receptionists are encyclopedias.  
 Some mobs are avalanches.  
 Some lies are boomerangs.  
 Some casinos are drugs.  
 Some schools are filters.  
 Some brains are warehouses.

---

## Appendix B

## Metaphors Used in Experiment 2

High similarity	Low similarity
Some arguments are wars.	Some conversations are wars.
Some lawyers are sponges.	Some teachers are sponges.
Some lies are boomerangs.	Some statements are boomerangs.
Some hippopotamuses are blimps.	Some lions are blimps.
Some horoscopes are maps.	Some books are maps.
Some saunas are ovens.	Some rooms are ovens.
Some ferries are bridges.	Some boats are bridges.
Some dragsters are rockets.	Some mopeds are rockets.
Some exams are filters.	Some applications are filters.
Some suburbs are parasites.	Some towns are parasites.
Some giraffes are skyscrapers.	Some busboys are skyscrapers.
Some auditions are doors.	Some plays are doors.
Some babies are angels.	Some plays are doors.
Some librarians are mice.	Some receptionists are mice.
Some stagecoaches are dinosaurs.	Some trains are dinosaurs.
Some salesman are bulldozers.	Some merchants are bulldozers.

Received September 22, 1997

Revision received July 29, 1999

Accepted August 12, 1999 ■