

Property Attribution in Metaphor Comprehension

SAM GLUCKSBERG

Princeton University

MATTHEW S. MCGLONE

Lafayette College

AND

DEANNA MANFREDI

National Analysts, Inc.

Metaphors of the form *X is a Y* may be conceived as statements of property attribution, in which properties of the vehicle *Y* are attributed to the topic *X*. The properties attributed from the vehicle to the topic are those that are epitomized by the vehicle, and may characterize a dimension of within-category variation in the topic. We report two experiments that tested the implications of this “property attribution” view of metaphor comprehension. One implication of this view is that metaphors are not implicitly transformed into comparisons (*X is like a Y*), but rather are understood as what they appear to be: class-inclusion assertions. If metaphors are understood as class-inclusion assertions, they should either change meaning or become nonsensical when reversed. Experiment 1 provides evidence for this claim. A second implication of the property attribution view is that topics and vehicles provide different kinds of information to guide and constrain the comprehension process. Experiment 2 provides evidence for this claim. These findings cannot be accounted for by models which assume that metaphors are understood as implicit comparisons. © 1997 Academic Press

The sentence forms *X is a Y* and *X is like a Y* are typically used to communicate very different relationships between the concepts *X* and *Y*. The first form is generally used to assert the membership of *X* in a category denoted by *Y*, as in *a lime is a citrus fruit* (see Brachman, 1983, for a discussion of the various uses of the ISA construction). The second form is used to assert that *X* and *Y* are similar in one or more respects, as in *a lime is like a*

lemon. Given the different conceptual relationships implied by each sentence form, it would seem that one cannot be used to paraphrase the other. The examples above conform to this rule. *A lime is like a lemon* cannot be acceptably paraphrased as *a lime is a lemon*. Similarly, *a lime is a citrus fruit* becomes anomalous when expressed as *a lime is like a citrus fruit*. However, the acceptability of paraphrasing metaphors as similes is a clear violation of this rule. For example, *rage is a volcano* does not change or lose meaning when paraphrased as *rage is like a volcano*.

Metaphor theorists have attributed the interchangeability of the category-inclusion and comparison forms to the procedures that people use to understand metaphors. The basic claim is that metaphors are first recognized as false categorical assertions, and then are implicitly transformed into comparison assertions. For example, Aristotle (1965) proposed

We are grateful for the financial support provided by the Public Health Service, Grant HD25826 to Princeton University, S. Glucksberg, Principal Investigator. The research reported here was performed as a Master's thesis and as a Doctoral dissertation submitted to Princeton University by Deanna Manfredi. We thank Philip Johnson-Laird, Dale Miller, and Steven Greene for their comments on an earlier draft of this manuscript. Address correspondence and reprint requests to Sam Glucksberg, Department of Psychology, Princeton University, Princeton, NJ 08544-1010. E-mail: samg@pucc.princeton.edu.

that all metaphors are fundamentally assertions of comparison, which may be expressed in category-inclusion form to add a rhetorical flourish of emphasis. Contemporary proponents of this comparison view have claimed that metaphor comprehension begins with the recognition of the comparison implied by a metaphor (Gentner, 1983; Miller, 1979; Ortony, 1979, 1993; Verbrugge & McCarrell, 1977). Once the implicit comparison is recognized, these theorists argue, it is then interpreted in the same manner as a literal comparison, by determining the relevant properties that the compared concepts have in common.

How are these common properties identified? The two most prominent comparison models of metaphor comprehension—Ortony's (1979) salience imbalance model and Gentner's (1983) structure-mapping model—assume that common properties are identified via a property-matching strategy. According to these models, metaphor comprehension begins with an initial exhaustive extraction of the properties comprising the representation of the topic and vehicle concepts. After the topic and vehicle properties have been extracted, they are exhaustively checked against one another, and those that "match" are taken as the grounds for the comparison.

Although elegant in its simplicity, the property-matching strategy posited by these models has limited explanatory value. As Ortony (1979) and Glucksberg and Keysar (1990) have pointed out, metaphoric comparisons that introduce new properties into an addressee's representation of a topic cannot be understood via a search for matching properties. For example, consider *Andrew's lecture was like a three-course dinner*. For people who are not familiar with the lecture (or lecturer) in question, there can be no a priori representation of the lecture that includes properties such as bountiful or sumptuous that can be matched to those of a three-course dinner. Yet these are exactly the sorts of properties that come to mind upon an uninformed reading of the statement. This argument applies with equal force to literal comparisons. If a person knows nothing about kumquats,

then telling her that *a kumquat is like an orange* will introduce new properties into her mental representation of kumquats (tangy, sweet, etc.), rather than produce a match between kumquat and orange properties.

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. Ortony (1979; Ortony et al., 1985) incorporates two assumptions into his salience imbalance model to account for the selection and application of vehicle properties to a topic in informative metaphoric comparisons. First, only those vehicle properties that exceed a salience criterion are considered for application to the topic. For the vehicle *three-course dinner*, properties such as bountiful and sumptuous (among others) would presumably be sufficiently salient to exceed this criterion. Second, matches between these candidate properties and those of the topic are determined using a similarity (as opposed to an identity) criterion. Thus, the bountiful amount and sumptuous quality of meals are similar if not identical to the bountiful amount and sumptuous quality of information that a lecture could contain, and consequently may be considered matches to these potential topic properties.

Glucksberg and Keysar (1990) frame the question of how informative metaphoric comparisons are understood in quite a different way. They begin by rejecting the notion that metaphors of the form *X is a Y* are implicitly converted into the simile form *X is like a Y* (for direct evidence against the implicit conversion claim, see Glucksberg, Gildea & Bookin, 1982). Instead, metaphors are understood as they are stated: categorization assertions in which the concept *X* is assigned to a category denoted *Y*. Thus, when someone says *my job is a jail*, the job in question is declared to belong to a category of unpleasant entities that is referred to by the word *jail*. This claim rests partly on the observation that many literal

comparisons are understood by casting the concepts being compared into a common category. For example, when people are asked how lemons and oranges are alike, the most usual answer is that they are both citrus fruits (Wechsler, 1958). When a comparison is intended to provide new information, as in *a dax is like an orange*, people may infer that the superordinate category exemplified by the vehicle, citrus fruits, constitutes the relevant ground of the comparison. Once the citrus fruit category is inferred, then properties of this category (e.g., pulpy flesh, tangy taste, Vitamin C content, etc.) can be attributed to the unfamiliar *dax*.

Metaphoric comparisons can be understood in essentially the same way. The statement *Andrew's lecture was like a three-course dinner* can be taken as an assertion that the lecture belongs to a category exemplified by three-course dinners. Of the numerous categories that three-course dinners could exemplify, only those that may plausibly contain lectures are ultimately considered for interpretational purposes. Thus, although three-course dinners can exemplify the category of expensive meals, this category is ultimately rejected because it cannot include lectures as members. On the other hand, three-course dinners can also exemplify a more abstract category of "things that come in large quantities and high quality," which can include lectures as members. Glucksberg and Keysar suggest that metaphoric comparisons are typically understood in terms of such abstract categories. When such a category is used to characterize a particular topic, it functions as an attributive category in that it provides properties (e.g., high quantity and quality, etc.) that may be attributed to the topic.

Although framed in different terms, Ortony's (1979) salience imbalance model and Glucksberg and Keysar's (1990) attributive categorization model offer similar accounts of the property attribution process in metaphor comprehension. However, there are two implications of the attributive categorization model which are inconsistent with Ortony's (1979) characterization of metaphors as implicit com-

parisons. First, if metaphors are indeed understood as class-inclusion assertions, then they should be nonreversible, just as literal class-inclusion assertions (e.g., *a robin is a bird*) are nonreversible (*a bird is a robin*). Although several studies have demonstrated that metaphors exhibit a higher degree of asymmetry than literal comparisons (Harwood & Verbrugge, 1977; Katz, 1982; Ortony et al., 1985), the stronger claim that metaphors are nonreversible has not been investigated. Second, the attributive categorization model implies that the topic and vehicle concepts make very different, albeit interactive, contributions to metaphor meaning. The vehicle provides properties to be attributed to the topic; one's familiarity with the topic provides constraints on the sorts of properties that may be plausibly attributed to it. In this sense, topics and vehicles play interactive roles in metaphor comprehension (Black, 1962, 1979).

We report two experiments that test the predictions of the attributive categorization model described above. If metaphors function as class-inclusion assertions, then they should not just be asymmetrical, but nonreversible, just as literal class-inclusion assertions are nonreversible. Experiment 1 assesses this claim. The claim that topics and vehicles play different roles implies that they provide different kinds of information to constrain and guide the comprehension process. Experiment 2 employed a priming paradigm introduced by Wolff and Gentner (1992) to assess this claim.

EXPERIMENT 1: METAPHORS ARE NONREVERSIBLE

The nonreversibility of metaphors and of metaphoric comparisons may manifest itself in at least two different ways. The clearest case of nonreversibility is displayed by statements that become anomalous when reversed. For example, the metaphor *alcohol is a crutch* becomes meaningless when reversed (*a crutch is alcohol*). Other metaphors display nonreversibility by virtue of conveying a completely different meaning when reversed. *My surgeon is a butcher* exemplifies this second case. In its original order, the statement can

be taken as a negative comment on the skill of some surgeons; when the topic and vehicle are reversed (*my butcher is a surgeon*), the statement becomes a positive comment regarding the skill of one's butcher. Because reversing the statement changes its meaning, this metaphor is also nonreversible. Note also that paraphrasing metaphors in comparison form does not weaken their nonreversibility: *alcohol is like a crutch* is also meaningless when reversed (*a crutch is like alcohol*); *my surgeon is like a butcher* also changes meaning when reversed as *my butcher is like a surgeon*.

In contrast to metaphoric comparisons, literal comparisons seem to be basically reversible. The assertion *yams are like potatoes* conveys the same meaning as *potatoes are like yams*: Both are starchy root vegetables with similar textures, tastes, and culinary uses. The asymmetry literal comparisons exhibit is attributable to conventions for marking given and new information, and/or a change in the degree of perceived similarity. In statements of the form *X is like a Y*, *X* and *Y* are typically the new and given information, respectively (Clark & Havilland, 1977). Because yams are generally less familiar than potatoes (at least to U.S. college students), people prefer to put yams in the new sentence position (the subject) and potatoes in the given sentence position (the predicate). With respect to degree of similarity, Tversky (1977) and others (e.g., Medin, Goldstone, & Gentner, 1993) have shown that reversing a literal comparison (and thus violating the given-new convention) can decrease the degree of perceived similarity. For example, people assign more credibility to the comparison *North Korea is like China* than to *China is like North Korea*. According to Tversky's contrast model, the properties that enter into the computation of similarity of the countries are identical for both statements. Differences in rated similarity are entirely accounted for by differential weighting of the distinctive properties of the subject and predicate terms of the comparison. The distinctive properties of the predicate are assigned more

weight than those of the subject (for details, see Tversky, 1977).

Property matching models can thus easily account for asymmetries in rated similarity of literal comparisons, as well as for ordering preferences (Medin et al., 1993). However, the sufficiency of this account for explaining metaphor reversal phenomena is questionable. When metaphors and similes are reversed, the effect is not simply a reduction in perceived similarity or ordering appropriateness; rather, the effect is that the meaning of the statement is lost or changed entirely. Thus, while *China is like North Korea* is interpretable (if not optimal), *a crutch is like alcohol* resists interpretation. *My butcher is like a surgeon* is interpretable, but is entirely different in meaning from *My surgeon is like a butcher*. These examples suggest that there is a qualitative difference between literal and metaphoric comparisons. How can this difference be explained?

One possibility is that metaphoric comparisons are more akin to class-inclusion assertions than they are to literal comparisons. As with metaphoric comparisons, the reversal of terms in a class-inclusion assertion (e.g., *a robin is a bird*) yields an anomalous statement (*a bird is a robin*). Glucksberg and Keysar (1990) suggest that the nonreversibility of metaphoric comparisons is a direct result of their being understood as class-inclusion assertions. A metaphor vehicle such as *crutch* is understood as referring to a higher-order category such as "things that people rely on in times of need" that may plausibly include the topic *alcohol* as a member. The comparison *alcohol is like a crutch* is comprehensible because the topic and vehicle terms may be understood as referring to their literal referents; however, the exemplar-category relationship between the topic and vehicle blocks a meaningful reversal of the statement (*a crutch is like alcohol*). When the topic and vehicle both have category referents, but different ones, as in *my surgeon is like a butcher* ("people who are skilled and precise" for *surgeon*, and "people who are sloppy and imprecise" for *butcher*), then the statement may be

reversed (*my butcher is like a surgeon*), but its meaning changes entirely.

These examples notwithstanding, the empirical evidence for this claim is equivocal. On the one hand, Gentner (1980) reported that reversing the topic and vehicle in a sample of eight metaphors had no discernible effects on people's judgments of metaphor aptness. However, the materials included metaphors such as the surgeon-butcher example described above (Ortony et al., 1985). Such statements may still be meaningful and apt when reversed, but the reversed meaning is entirely different from the original meaning. Metaphors may also maintain meaningfulness when the reversal is perceived as an instance of poetic inversion, as in *a mighty fortress is our God*. In such cases, the reversal is perceived as a surface grammatical inversion of the subject and predicate, and is therefore implicitly reversed (*our God is a mighty fortress*) during comprehension.

Ortony et al. (1985) dealt with the problem of implicit reversals by giving subjects pairs of literal comparisons and metaphoric comparisons, with the order of subject and predicate reversed in each pair. Subjects were asked to decide which order in each pair "was more reasonable, sensible, appropriate, etc." (p. 575) and then to rate the similarity of the terms in each comparison, with the preferred order rated first. As expected on the salience imbalance model, reversed metaphoric comparisons were rated as less reasonable, sensible, and appropriate than reversed literal comparisons. Also as expected, the terms of literal and metaphoric comparisons were rated as equally similar when in original order, but the terms of metaphoric comparisons were rated as less similar when in reversed order. These data support the claim that metaphoric comparisons are asymmetrical, but do not speak directly to the stronger claim that metaphoric comparisons are nonreversible. To support this latter claim, we need to demonstrate that reversed metaphors are not simply less reasonable, but that they are either anomalous or interpreted in new ways. Experiment 1 provides such a demonstration.

Method

Subjects. Forty Princeton undergraduates, all native English speakers, served as paid volunteers. None had participated in any prior studies of figurative language.

Materials and procedure. Twenty-four metaphors and their corresponding similes, and 12 literal similarity statements were created for this experiment. These materials are presented in Appendix A. For each of the original-order statements, two reversed order statements were generated. One statement simply interchanged the topic and vehicle nouns; the other interchanged the topic and vehicle noun phrases. For example, the original-order metaphor *my marriage was an icebox* was reversed to produce *my icebox was a marriage* and *an icebox was my marriage*. Both of these reversal types were incorporated into our materials because we could not make any a priori decisions as to which reversal type would be more comprehensible or pragmatically acceptable. Because we predicted that metaphors and metaphoric comparisons would be nonreversible, we used the conservative strategy of including only the reversal order of each pair with the higher comprehensibility rating for the analyses that we report below.

The original statements combined with the two types of reversals yielded nine experimental sentence types: metaphors, metaphoric comparisons, and literal comparisons in original, noun-reversed, and noun-phrase reversed order. These materials were partitioned to create eight different questionnaire booklets so that each subject saw any given subject-predicate combination only once: 12 each of metaphors, metaphoric comparisons and literal comparisons. Filler items remained constant for each booklet, as did seven warm-up items to familiarize subjects with the rating scales and paraphrase instructions. The order of items was randomized across the eight different versions of the questionnaire. Each version was completed by five subjects. Subjects were seated individually in a small room and instructed to read and rate each statement on a scale of zero to seven, with zero indicating

that the statement made no sense at all, and seven indicating that the statement made perfect sense. For each statement that was rated one or higher, subjects were asked to write a paraphrase of the statement to indicate what a speaker would intend to communicate by using that statement. The paraphrases for the metaphors, metaphoric comparisons and literal comparisons were then submitted to two independent judges for classification. Because comparisons, especially metaphoric comparisons, could often be identified as either in preferred or reversed order, the judges could not be blind to experimental condition. With this caveat, the paraphrases of the reversed statements were classified as follows. We first examined the paraphrases of the original-order statements. As expected there was virtually unanimous agreement as to the grounds of the original-order metaphors, metaphoric comparisons and literal comparisons. These paraphrases served as the basis for assessing the paraphrases of the reversed statements. The two judges classified the paraphrases in two stages. In the first stage, each paraphrase was judged to be either acceptable or unacceptable. The acceptable category included all paraphrases of reversed statements that were judged to be equivalent in meaning to the original-order statement, and whose meaningfulness rating was within the 95% confidence interval of the original-order rating. Those reversed statements that were judged to be different in meaning—i.e., different ground for the comparison or metaphor—were placed in the unacceptable category, and then further classified into four subcategories, as follows:

- a. *Uninterpretable*: Rated meaningfulness was zero and hence no paraphrase given.
- b. *Less meaningful*: The meaningfulness rating for the reversal was significantly lower than the rating for the original order (exceeded 95% confidence interval).
- c. *Re-reversals*: The meaningfulness rating was not significantly different, but the paraphrase indicated that the statement had been transformed back into the original order, either explicitly or implicitly. For example, in re-

sponse to the reversed simile *Sharks were like his lawyers*, one subject explicitly reversed the terms in the paraphrase “His lawyers were like sharks.” Other paraphrases were classified as “implicit” reversals when they involved the attribution of properties from the canonical topic to the canonical vehicle. For example, in response to the reversed metaphor *Her ladder was an education*, one subject generated the paraphrase “She climbed up the various levels of schooling to reach her goal at the top.” Thus, schooling is characterized as something to be climbed (i.e., as a ladder), rather than a ladder being characterized as an education of some sort.

d. *New ground*: The paraphrase of the reversal specified a different ground for the comparison. For example, paraphrasing the original metaphor *large billboards are warts* as “large billboards scar the face of the landscape as warts scar the human body” indicates that the property “ugliness” is part of the ground. The paraphrase “large warts are highly visible” for the reversed metaphor *Large warts are billboards* expresses a different ground, namely, the high visibility of billboards.

Examples of each of these types of paraphrases are shown in Table 1, together with their meaningfulness ratings. There was 90% agreement between the judges’ initial independent classifications. The remaining differences were resolved in discussion.

Results and Discussion

Although we collected meaningfulness ratings, these were intended primarily for use in classifying the paraphrases for the three sentence types of interest: metaphors, similes, and literal comparisons. The meaningfulness ratings are not particularly informative because reversed metaphors and similes that were either explicitly or implicitly re-reversed could (and often did) receive high meaningfulness ratings. Nonetheless, even though they may seriously underestimate the degree of nonreversibility of metaphors, the mean ratings do indicate that literal comparisons are far more

TABLE 1

EXAMPLES OF PARAPHRASE CLASSIFICATIONS (MEANINGFULNESS RATINGS IN PARENTHESES)

A. Acceptable reversals	
Original order:	His Ph.D. is like an M.D. (6), paraphrased as: Same amount of years in school.
Reversal:	His M.D. is like a Ph.D. (6), paraphrased as: Both are advanced terminal degrees.
Original order:	Her cigarettes are pacifiers (7), paraphrased as: Cigarettes calm her.
Reversal:	Her pacifiers are cigarettes (5), paraphrased as: Both are relaxing.
B. Unacceptable reversals	
Original order:	Their trust is glue (7), paraphrased as: Their trust is as lasting and bonding as glue.
Reversals	
a. Uninterpretable:	Glue is their trust (0), no paraphrase provided.
b. Less meaningful:	Glue is their trust (2), paraphrased as: They trust the glue to work.
c. Re-reversals	
1. Explicit re-reversals:	Their glue is trust (2), paraphrased as: Rather their trust is like glue-never ending.
2. Implicit re-reversal:	Glue is like trust (3), paraphrased as: Their trust holds them together like glue.
d. New Ground	
1. Original order:	Large billboards are warts (6), paraphrased as: Large billboards scar the face of the landscape as warts scar the human body.
2. Reversal:	Large warts are billboards (5), paraphrased as: Large warts are highly visible.

reversible than are metaphors. Table 2 presents the mean ratings for metaphors and literal comparisons in both their original and best reversal (i.e., the reversal that received the highest meaningfulness ratings) orders.

Initial analyses indicated that booklet version did not produce any significant main effects or interactions with other independent variables, and so further analyses were collapsed across this variable. Two additional ANOVAs were conducted with statement type (metaphor, simile, and literal comparison) and order (original order and best reversal) as within-subject factors, one with subjects as a random factor (F_s) and one with items as a random factor (F_i). These analyses revealed a reliable main effect for order ($F_s(1,39) = 6.89, p < .01; F_i(1,35) = 5.40, p < .05$) and a

reliable statement type \times order interaction ($F_s(2,78) = 5.15, p < .05; F_i(2,70) = 4.76, p < .01$). Pairwise contrasts revealed that reversed metaphors received significantly lower meaningfulness ratings than their original-order counterparts ($t_s(39) = 3.49, p < .01; t_i(35) = 2.44, p < .02$) and the same was true for reversed similes ($t_s(39) = 2.92, p < .01; t_i(35) = 2.12, p < .05$). However, literal comparisons did not differ in meaningfulness as a function of order ($t_s(39) = 1.17, p > .05; t_i(35) < 1$). Clearly, metaphors and similes behave alike, and both differ significantly from literal comparisons. The metaphoric statements are less meaningful when reversed, while the literal comparisons remain relatively unaffected.

The more informative data stem from the paraphrase analysis. For this analysis, only the

TABLE 2

MEAN MEANINGFULNESS RATINGS AS A FUNCTION OF STATEMENT TYPE AND TERM ORDER
(STANDARD ERRORS IN PARENTHESES)

Term order	Metaphors	Metaphoric comparisons	Literal comparisons
Original order	5.7 (1.6)	5.5 (1.9)	4.2 (2.4)
Best reversal	3.1 (1.5)	3.4 (1.7)	4.9 (1.2)

higher-rated reversals were included so as to avoid cases in which reversals were unacceptable because of the characteristics of particular determiners or possessives. For example, the original-order literal comparison *his dog is like a cat* was judged acceptable only 60% of the time when reversed as *a cat is like his dog*, but was judged acceptable 100% of the time when reversed as *his cat is like a dog*. Only paraphrases for the latter member of this pair were included in the analysis. The results of the paraphrase analysis clearly indicate that metaphors and literal statements differ with respect to their reversability. Less than 4% of metaphoric statements were acceptable when reversed, compared to 82% of the literal comparisons. Conversely, only 18% of literal statements were unacceptable when reversed, compared with 81 and 78% for metaphors and similes, respectively. Of the unacceptable reversals, 15.5% of the metaphors and 22.3% of the similes were explicitly transformed back into their original orders. Fifty-five percent of the metaphors and 33% of the similes were implicitly re-reversed into their original orders. Finally, 18% of reversed metaphors and 18% of the similes were interpreted with new grounds, while none of the reversed literal comparisons were interpreted in this manner. Combining unacceptable reversals and new grounds for the metaphors and similes, virtually all (97.5%) were judged to be unacceptable in reversed form, compared to only 18% for the literal comparisons.

These data provide strong support for the claim that metaphors and similes either lose or change meaning when reversed. Literal comparisons, in contrast, are generally reversible. Several literal comparisons were judged to be less meaningful when reversed, but none of the reversals was interpreted as implying a ground different from that of the original order. This finding is consistent with Tversky's (1977) claim that reversed literal comparisons may differ in acceptability without any change in the perceived ground of comparison.

The observed difference in ratings of meaningfulness between reversed similes and literal comparisons replicates the results of Ortony

et al. (1985, Experiment 1). These researchers explained this finding in terms of differences in property salience between the concepts being compared. According to their proposal, literal comparisons involve matches of properties that are highly salient for both the subject and predicate. For example, the property "shiny" is highly salient for both terms in the comparison *copper is like brass*. In contrast, canonical similes involve matches of properties that are highly salient for the predicate (vehicle), but not for the subject (topic). Thus, the property "vicious" is highly salient for the vehicle but not the topic of the assertion *her words were like daggers*. When the terms of a simile are placed in noncanonical (i.e., reversed) order, the interpreter's expectations are violated, and consequently the non-canonical simile is perceived as less meaningful.

While our results are consistent with this proposal, we argue that the characterization of reversed similes (and reversed metaphors) as less meaningful is inappropriate. Our results suggest that reversed metaphoric statements are not perceived as merely less meaningful, but instead as either not meaningful at all or meaningful in a completely different way. In this respect, metaphoric statements such as *her words were (like) daggers* behave as class-inclusion assertions, in that reversing the terms (*her daggers were (like) words*) yields an anomalous statement. This suggests that metaphoric comparison statements convey a subordinate-superordinate relationship that literal comparisons clearly do not. This relationship is conveyed, Glucksberg and Keysar (1990) argue, because the vehicle term *daggers* may be understood as referring to a higher-order category (e.g., "things that cause suffering").

EXPERIMENT 2: TOPIC AND VEHICLE INFORMATIVENESS

The claim that metaphoric statements are understood as class-inclusion assertions implies that the topic and vehicle make very different, although interactive, contributions to metaphor meaning. The vehicle provides candidate properties that may be attributed to the

topic. In contrast, the topic provides constraints on the sorts of properties that may be plausibly attributed to it (Ortony's 1979 salience imbalance model makes a similar claim). This claim raises two important questions. First, what types of constraints does the topic impose on the property attribution process? Second, how are candidate properties of the vehicle identified and selected?

For any given metaphor topic, only certain sorts of property attributions will be interesting and/or relevant. The relevance of a given property to a topic can best be described at the level of dimensions. When the topic is *road*, for example, dimensions such as shape (e.g., straight, curved, twisting), surface (smooth or bumpy), and width (narrow or wide), are meaningful and relevant in most contexts in which roads are discussed. Dimensions such as cost (cheap, expensive) and color (black, white, gray) can be meaningful, but are irrelevant in most contexts in which roads are discussed. Still other dimensions, such as emotional arousal (calm, neutral, excited) are not applicable to roads, and consequently characterizations on these dimensions would be meaningless. Another way to specify those dimensions that would be meaningful and relevant for a metaphor topic would be to specify the dimensions of within-category variation. For a concept such as *roads*, the ways that roads can meaningfully differ from one another constitute that concept's relevant dimensions for attribution.

Our notion of relevant dimensions for attribution is analogous to the notion of relevance in the conceptual combination literature (e.g., Murphy, 1988, 1990). When asked to interpret adjective–noun or noun–noun combinations such as *blind lawyer* or *jail-job*, existential possibility is often less important than plausibility. For example, people have difficulty interpreting the combination *unframed planet*, even though they will readily agree that planets are never framed. This noun phrase is difficult to interpret because the adjective *unframed* does not characterize a dimension that is relevant for the concept *planet*. In terms of within-category variation, the “framing”

dimension does not distinguish between members of the category “planets.” Other nouns, such as *photograph*, or *painting*, may plausibly differ on this dimension, and thus the combinations *framed photograph* and *unframed painting* are interpretable. We argue that metaphor topics behave as do the head nouns in conceptual combinations, in that characterizations of topics are meaningful only when they are made along relevant attributional dimensions.

The number of relevant attributional dimensions will vary from topic to topic. Topics with relatively few such dimensions will place a high level of constraint on potential attributions. The topic *lawyer*, for example, is likely to be characterized on relatively few dimensions, among them skill, experience, temperament, ambition, reputation and cost. It would be highly unlikely that any given lawyer would be characterized *qua* lawyer on dimensions that are irrelevant to the practice of law, such as height, weight, or musical ability. Topics such as *lawyer* thus impose a high level of constraint on potential attributions. In contrast, other topics such as *my brother* provide very few constraints on potential attributions because one might say almost anything about one's brother. Metaphor topics, then, can vary in terms of the level of constraint that they place on interpretation. High-constraint topics produce limited expectations about how they might be characterized. Low-constraint topics produce relatively unlimited expectations about how they might be characterized.

Just as topics can vary in the number of relevant attributional dimensions, metaphor vehicles can vary with respect to the number or variety of properties they can provide as candidate attributes. Some metaphor vehicles are unambiguous in the sense that they uniquely exemplify an attributive category. When used as vehicles, terms such as *shark* and *jail* are relatively unambiguous in this way. *Shark* is emblematic of the category of vicious predators; *jail* is emblematic of situations that are unpleasant and confining. Other metaphor vehicles are relatively ambiguous because they do not uniquely exemplify an

attributive category. For example, consider *our love is a voyage to the bottom of the sea*. Because such a voyage does not exemplify any category in particular, it is unclear what properties this vehicle might provide to characterize a metaphor topic.

Understanding a metaphor thus requires two kinds of world knowledge. First, one must know enough about the topic to appreciate which kinds of characterizations are relevant and meaningful (i.e., the relevant dimensions of within-category variation of the topic concept). Second, one must know enough about the metaphor vehicle to know what kinds of things it may epitomize. Given this kind of knowledge base, metaphors with ambiguous vehicles can be readily understood when the metaphor topic is high-constraint (i.e., has relatively few attributional dimensions). Similarly, metaphors with low-constraint topics (i.e., with many attributional dimensions) can be readily understood when the metaphor vehicle is reasonably unambiguous. In this sense, metaphor topics and vehicles are used interactively to generate interpretations. The interactive property-attribution view we propose here is similar to that described by the philosopher Max Black (1962, 1979).

One empirical implication of this view is that metaphor topics and vehicles contribute different kinds of information to guide comprehension. One way to test this implication was suggested by Wolff and Gentner (1992), who used a priming paradigm to investigate the different roles of the topic and vehicle in immediate metaphor comprehension. Their reasoning was straightforward: if metaphor topics and vehicles each provide useful information for metaphor comprehension, then providing either the topic or the vehicle in advance of the metaphor itself should facilitate comprehension. Topics and vehicles were each used as primes in a timed metaphor comprehension task. The results suggested that both topics and vehicles facilitated processing, relative to a no-prime control condition. Wolff and Gentner interpreted this finding as supporting a property-matching account of metaphor comprehension. They argued that either

the topic or vehicle provided in advance of the metaphor itself permitted the reader to get a head start on the exhaustive extraction of properties that is required before property matching can begin.

The property attribution view we propose suggests that their findings may reflect an adventitious choice of materials, rather than a property extraction/matching process. Recall that topics can vary in attributional constraint. High-constraint topics have few attributional dimensions, and thus presenting such topics in advance of a metaphor may facilitate comprehension. In contrast, low-constraint topics have many such dimensions, and so provide little useful information when presented in isolation. Accordingly, high-constraint topics should be effective primes for metaphor understanding, while low-constraint topics should be less so. On the property matching account Wolff and Gentner describe, the facilitation produced by advance presentation of the topic should not depend on the attributional constraints imposed by the topic. The degree of vehicle ambiguity presents a similar potential problem. Unambiguous vehicles, when presented in isolation, should provide useful information for subsequent metaphor processing, while ambiguous vehicles should not. Again, Wolff and Gentner did not make this distinction when they argued that the advance presentation of any metaphor vehicle should facilitate processing because any vehicle would permit a head start on property extraction.

Because Wolff and Gentner did not distinguish between low- and high-constraint topics or between ambiguous and unambiguous vehicles, their results cannot distinguish between a property extraction/matching account and a property attribution account of metaphor comprehension. In Experiment 2, we employed a variant of the Wolff & Gentner priming paradigm to investigate the predictions made by the property attribution model described above. We expected that high-constraint topics would be effective as primes for metaphor comprehension, while low-constraint topics would be either less effective or not effective

at all. Analogously, we expected that unambiguous vehicles would be effective primes, while ambiguous vehicles would not.

Method

Subjects. Seventeen undergraduate students at Princeton University, all native English speakers, served as paid volunteers. None had participated in any previous studies of figurative language comprehension. An additional 20 participants from the same student sample provided ratings for the materials, as described below.

Materials. The experimental sentences were constructed in two stages. First, four separate lists of candidate terms were generated: potentially low-constraint topics, high-constraint topics, ambiguous vehicles, and unambiguous vehicles. We chose candidate terms that intuitively seemed to meet the criteria for each category: low-constraint topics are those that can plausibly be described in many meaningful ways, high-constraint topics are those that can be described in relatively few ways; unambiguous vehicles are those about which there is high agreement concerning their attributional properties, and ambiguous vehicles are those for which there is low such agreement. We conducted a pretest on a large pool of metaphor stimuli to identify topic and vehicle terms that met our design constraints. The experimental topics and vehicles chosen from this pool are presented in Appendix B.

Up until this point in the materials construction, we had assumed that level of constraint (for topics) and ambiguity (for vehicles) were orthogonal dimensions; however, we had yet to empirically demonstrate that they were. Demonstrating the orthogonality of these dimensions is necessary to establish that the topic and vehicle make role-specific contributions to metaphor meaning. To this end, we conducted two manipulation checks to assess the level of constraint and ambiguity of our topic and vehicle stimuli.

For the first check, a questionnaire was constructed containing the 48 experimental topic and vehicle terms in a random order. The instructions for the questionnaire asked subjects to list for each term as many questions as

they could think of that would distinguish one instance of the term concept from another. The number of questions generated for each term served as our operational definition for level of constraint: terms for which there were few questions generated were high-constraint, and terms for which many questions were generated were low-constraint. Ten Princeton undergraduates (none of whom participated in the experiment proper) completed this questionnaire. The results offer empirical support for our intuitive judgments of topic level of constraint: subjects generated on average 3.4 ($SD = .4$) and 6.2 ($SD = 1.0$) questions for the topics we had chosen as high-constraint and low-constraint, respectively. However, there was no appreciable difference in the mean number of questions generated for the vehicles: on average, unambiguous vehicles elicited 4.2 questions ($SD = .5$) and ambiguous vehicles elicited 4.8 questions ($SD = .4$). This latter finding indicates that in choosing our vehicle terms, we did not inadvertently choose terms that differed significantly in level of constraint.

For the second manipulation check, a questionnaire was constructed in which each of the 48 topic and vehicle terms were presented in the sentence frame *X is a [term]*. Subjects were instructed to treat each sentence frame as a metaphor in which the topic had been deleted, and to generate a property of the deleted topic X implied by the metaphoric assertion. A second set of 10 Princeton undergraduates (none of whom participated in the experiment proper) completed this questionnaire. The percentage of subjects who generated the same property for a given term served as our operational definition of ambiguity. For the unambiguous and ambiguous vehicles, subjects generated the same property 58.5% ($SD = 6.0\%$) and 21.8% ($SD = 4.2\%$) of the time, respectively. For the high- and low-constraint topics, subjects generated the same property 27.7% ($SD = 4.4\%$) and 28.3% ($SD = 4.7\%$) of the time, respectively. The fact that the low- and high-constraint topics did not appreciably differ in ambiguity complements our earlier finding regarding level of constraint for vehi-

cles, indicating that level of constraint and ambiguity are orthogonal dimensions (at least for the terms used in our experiment).

The two types of metaphor topics and vehicles were used to generate four kinds of metaphors for the experiment proper. For each of the 24 low- and high-constraint topics, a vehicle was chosen that, in the experimenters' judgment, was relatively unambiguous and produced a sensible metaphor. Similarly, for each of the 24 unambiguous and ambiguous vehicles, a topic was chosen that, again in our judgment, was moderately constraining and also produced a sensible metaphor. Our intent was to generate a set of four kinds of metaphors that would be equally comprehensible when presented in isolation. These 48 experimental metaphors are presented in Appendix B.

Primes were constructed for each metaphor by pairing the experimental topic or vehicle with a row of asterisks in the metaphor sentence form *X is a Y* (or *Xs are Ys*). For the metaphors that differed in level of topic constraint, the vehicle term was omitted (e.g., *Some jobs are ******); for metaphors that differed in vehicle ambiguity, the topic was omitted (e.g., *Some ***** are viruses*). Each metaphor was preceded by one of these prime types (prime condition), or by no prime at all (no-prime condition). Stimulus lists were arranged so that all metaphors appeared in the prime and no-prime conditions equally often. Item ordering and list composition were controlled by Micro-Experimental Laboratory (MEL) programming, and each subject viewed the items in a different random order.

Design and procedure. This experiment employed a 2×2 (term type \times prime informativeness) within-subjects design. Upon arrival at the laboratory, subjects were told that the study was an investigation of people's ability to understand figurative language. Each subject was then invited into a sound-attenuated booth and seated in front of a computer screen, where the following instructions appeared:

In this study you will be asked to read a series of metaphors. Metaphors are figurative statements such as Shakespeare's *All the world's a stage* or the common expression *Some lawyers are sharks*.

On some of the trials, before seeing the complete metaphor you will see part of it for two seconds. When you are presented with this prior information about the metaphor, please read it and think about the ways that information might be used to make a sensible metaphor. On some trials, you will not be given any prior information about the metaphor. On these trials you should just focus your attention on the row of asterisks that will appear in the center of the screen and wait for the metaphor to be presented.

On either trial type, your job is to simply read the metaphor and strike the space bar on the keyboard in front of you when you feel you have understood it. Once you strike the space bar, the metaphor will be removed from the screen and a new trial will begin. After you have completed the experiment, you will be presented with a brief questionnaire containing some of the metaphors you read in the study. Certain words will be left out of each metaphor and you will be asked to replace them to the best of your memory. It is very important that you carefully read all the information presented during the experiment in order to do well on the memory test.

After the subject read the instructions, any questions s/he asked were answered, and then three practice trials were presented, followed by the experiment proper. A trial presentation consisted of a row of asterisks (e.g., *******) presented for 500 ms, a priming string (e.g., *Some jobs are ******) presented for 2 s (on no-prime trials, the row of asterisks remained on the screen), and then the target sentence (e.g., *Some jobs are prisons*) presented until a key press. Response timing was initiated when the target metaphor appeared on the screen and stopped when the space bar was pressed.

When subjects completed the experiment, they left the booth and were seated at a desk to complete a questionnaire consisting of 10 of the metaphors that they had just seen with either the topic or vehicle omitted (5 of each kind). Subjects were asked to complete each item as they remembered it, and any subject who failed to complete at least 6 of the 10 items correctly were dropped (i.e., the data were discarded) and replaced.

Results and Discussion

The data of two subjects were discarded because of their poor performance in the ques-

TABLE 3

MEAN RESPONSE TIMES (IN MS) AS A FUNCTION OF TERM TYPE, TERM INFORMATIVENESS, AND PRIMING CONDITION

Term type	Informativeness	No-prime condition	Prime condition
Topic	High-constraint	2261	1445
	Low-constraint	2321	2143
Vehicle	Unambiguous	2284	1292
	Ambiguous	2302	2330

tionnaire post-test and were replaced by data from two additional subjects to provide 15 subjects' data for analysis. The mean comprehension times for the metaphor stimuli in the no-prime and prime conditions are presented in Table 3. Priming effects were assessed by computing the mean difference scores between the no-prime and prime condition comprehension times for each metaphor type. These mean difference scores were then submitted to 2×2 analyses of variance, with term type (topic or vehicle) and prime informativeness (informative or uninformative) as repeated factors. Two such ANOVAs were conducted: one with subjects as a random factor (F_s), and one with items as a random factor (F_i). For these analyses, high-constraint topics and unambiguous vehicles were classified as informative primes; low-constraint topics and ambiguous vehicles were classified as uninformative.

As expected, there was no main effect of term type ($F_s(1,14) = 1.01, p > .05$; $F_i(1,46) = 1.41, p > .05$). Also as expected, informative primes facilitated comprehension and uninformative primes did not, as reflected by the main effect of prime informativeness ($F_s(1,14) = 24.91, p < .01$; $F_i(1,46) = 26.63, p < .01$). Primes consisting of high-constraint topics or unambiguous vehicles yielded priming effects of 816 and 992 ms, respectively. Primes consisting of low-constraint topics or ambiguous vehicles yielded priming effects of 178 and -28 ms, respectively. These data are consistent with the claims that (a) level of constraint is an important characteristic of metaphor topics, and (b) ambiguity is an important characteristic of metaphor vehicles.

However, the role-specificity hypothesis entails two additional claims: that level of constraint is not an important characteristic of metaphor vehicles, and that ambiguity is not an important characteristic of metaphor topics. In the manipulation check, we found that high- and low-constraint topics did not vary in ambiguity and that unambiguous and ambiguous vehicles did not vary in level of constraint. While these findings offer some support for the role-specificity hypothesis, a more stringent test is possible. Because we had obtained ratings of constraint levels and ambiguity for all of our experimental items, we could assess whether or the observed priming effects were mediated by topic ambiguity and/or vehicle constraint level.

In order to address this issue we used the topic ambiguity rating data from the manipulation check to divide both low- and high-constraint topic conditions into two types: ambiguous and unambiguous topics via a median split. Topics considered unambiguous elicited a mean agreement level of 56%, comparable to the agreement level of 58% for the unambiguous vehicles used in this study. Topics considered ambiguous elicited an agreement level of 35%, compared to the 22% level for the ambiguous vehicles. Mean priming effects as a function of topic constraint and ambiguity are provided in Table 4. Analyses of variance applied to these data yielded a reliable effect of topic constraint ($F_s(1,14) = 9.46, p < .01$; $F_i(1,23) = 8.89, p < .01$). There was no effect of topic ambiguity ($F_s(1,14) = 2.15, p > .05$; $F_i(1,23) = 2.00, p > .05$), and no interaction between topic constraint and ambiguity ($F_s(1,14) = 1.22, p > .05$; $F_i(1,23) = 1.91, p > .05$).

TABLE 4

MEAN PRIMING EFFECT (MS) AS A FUNCTION OF TOPIC AND VEHICLE CHARACTERISTICS

Term type	Constraint level	Level of ambiguity	
		Ambiguous	Unambiguous
Topic	High	886	746
	Low	238	118
Vehicle	High	-7	1058
	Low	-49	926

An analogous analysis of vehicle constraint and ambiguity yielded similar results. Low- and high-constraint vehicles had constraint levels of 3.8 and 4.9, respectively. The original low- and high-constraint levels were 3.2 and 6.2, respectively. Priming effects as a function of vehicle ambiguity and level of constraint are presented in Table 4. A 2×2 analyses of variance applied to these data yielded a reliable effect of vehicle ambiguity ($F_s(1,14) = 11.07, p < .01$; $F_i(1,23) = 13.32, p < .01$), but no effect of level of constraint ($F_s(1,14) = 1.74, p > .05$; $F_i(1,23) = 1.84, p > .05$). The interaction between ambiguity and level of constraint was not significant on either analysis (F_s and $F_i < 1$). It could, perhaps, be argued that the sample of items that we used did not vary sufficiently in terms of ambiguity or level of constraint to provide a sufficiently sensitive test of the hypotheses. Within the context of this experiment, however, topic ambiguity and vehicle level of constraint had no discernible effects.

The pattern of priming results obtained in this experiment is inconsistent with the property extraction/matching model of metaphor comprehension described by Wolff and Gentner (1992). According to these researchers, advance presentation of *any* metaphoric term should expedite the property extraction process, and thereby facilitate metaphor comprehension. However, we found that only metaphoric terms with particular characteristics produced such facilitation. Specifically, high-constraint topics and unambiguous vehicles produced facilitation, while low-constraint

topics and ambiguous vehicles did not. The relationship between term characteristics and prime effectiveness, we argue, reflects the role-specific contributions that topics and vehicles made to metaphor meaning overall. High-constraint topics generated fewer expectations than low-constraint topics regarding the dimensions upon which they were likely to be characterized by a vehicle; analogously, unambiguous vehicles generated fewer expectations than ambiguous vehicles regarding the properties they provided for attribution to the topic. As a result, high-constraint topics and unambiguous vehicles offered more advance information when presented as primes than their low-constraint and ambiguous counterparts.

GENERAL DISCUSSION

In perhaps the most articulated development of the comparison view, Miller (1979) argued that “. . . *Man is a wolf* is false in fact. In order to understand it, the reader must associate it with *Man is like a wolf*, or, even weaker, *Man seems like a wolf* (to the author)” (p. 214). Miller concludes that “the grounds for a metaphor . . . can be formulated as relations of similitude that can be expressed as comparison statements” (p. 248).

Previous studies have found little support for Miller's claims. To begin with, the comparison view implies that metaphor processing is not only slower than literal comprehension but optional: a metaphor must first be recognized as false and rejected before it is treated as a relation of similitude. Neither of these implications is supported by the available evidence. Ortony and others (Ortony, Schallert, Reynolds, & Antos, 1978; Harris, 1976) have found that metaphors are processed as quickly as comparable literal expressions, casting doubt on the notion that metaphor processing is slower than literal processing. With respect to the priority of literal interpretation, Glucksberg and others (Gibbs, 1984; Glucksberg, Gildea, & Bookin, 1982; Keysar, 1989) have shown that false literal meanings need not be rejected before metaphorical meanings are processed.

The experiments we report here cast further doubt on the comparison view. First, we found that metaphors and similes do not behave as do literal comparisons (Experiment 1). While literal comparisons were perceived as reversible by our subjects, both metaphors and similes were perceived as nonreversible. In this respect, metaphoric statements behave as do literal class-inclusion assertions, not as literal comparisons. Second, the pattern of priming effects found in Experiment 2 does not support a property extraction/matching model of the sort described by Wolff and Gentner (1992). If metaphoric statements were understood by a two-stage property extraction and matching process, then prior presentation of either a topic or a vehicle should allow the reader to get a head start on the property extraction stage. We found that prior presentation of a topic or a vehicle facilitated metaphor comprehension only when the topic was high-constraint or the vehicle was unambiguous. When the topic was low-constraint or the vehicle ambiguous, there was no facilitation of metaphor comprehension.

The overall pattern of priming is thus most compatible with an interactive property attribution model of metaphor comprehension. More telling, this pattern is a natural outcome of that model. Neither Gentner's (1983) structure-mapping model nor Ortony's (1979) salience imbalance model provides a rationale for (a) classifying metaphor topics and vehicles in terms of constraint and ambiguity levels, or (b) for predicting their differential effectiveness as primes for metaphor comprehension. At the very least, matching models of metaphor comprehension will need to be elaborated to include the discourse-relevant constructs of topic constraint levels and vehicle ambiguity in addition to the structural level constructs of property listings and property salience.

How might matching models be elaborated? As currently formulated, matching models such as Tversky's (1977) contrast model, Gentner's (1983) structure mapping model, and Ortony's (1979) salience imbalance model are more applicable to the task of judg-

ing degree of similarity between two concepts than to the task of understanding similarity assertions in the first place. When two concepts X and Y are compared, the greater the number of properties they share, the greater the degree of judged similarity. However, as Tversky (1977) noted, people's interpretations of comparison statements make it clear that not all the properties of X and Y are considered. For example, when people interpret the statement *Yale is like Harvard*, they consider commonalities such as reputation and cost, while excluding such commonalities as the presence of libraries, dormitories, deans, etc. As Goodman (1972) noted, the number of properties that any two concepts may share is unlimited. Thus, any property-matching model of similarity must postulate some mechanism for property selection. How might properties be selected? One possibility is that properties of the comparison predicate are selected that are relevant and meaningful when applied to dimensions of the subject. Such a selection process can account for the properties that are considered when literal comparisons are understood, and may also be applied to metaphors and similes.

Consider the metaphor vehicle *gold mine*. This concept includes such concrete properties as containing a precious metal, being a source of material wealth, and so forth. These properties are specific instantiations of a more general concept that, for purposes of illustration, we will refer to as "things of high value." This more general property is available to characterize any topic that may vary on the value dimension. If this dimension is relevant to a topic, then that topic may be meaningfully characterized as a gold mine. The specific manner in which high value is instantiated will differ from topic to topic. Thus, a library may be a gold mine in that a library contains a wealth of information, while an invention can be a gold mine in that it can earn the inventor a great deal of money.

A modification of property-matching models along these lines could, in principle, account for comparisons in which the two concepts being compared are relatively familiar

to the reader. Such models would still be problematic, however, for comparisons in which the reader is relatively unfamiliar with the subject. For metaphors and similes, this occurs whenever a topic has no preexisting properties in common with the vehicle. If such statements were rare, they would not pose a serious problem. However, they may be the rule rather than the exception. In many situations, an addressee may be familiar only with the general properties of a topic. For example, consider *Deanna's job is like a prison*. If an addressee is not familiar with the particular job in question, then only the general properties of jobs are available. The properties that differentiate this job from others, such as high- vs low-paying, interesting vs boring, etc. are not available. Yet these are precisely the kinds of properties that informative assertions are used to convey. In ordinary discourse, people rarely communicate definitional properties, as in *jobs are paid activities*. Presumably, people within a speech community presuppose common knowledge and common ground about the core properties of concepts. Instead, informative assertions convey new information, such as the confining or unpleasant nature of a specific job, in contrast to other jobs that may not be.

Informative metaphoric assertions thus do not, in general, involve a search for matching properties between a topic and a vehicle. If an addressee's mental representation of a topic and vehicle of such an assertion included the properties that they have in common, then the assertion would not be informative! Informative comparisons in general, and metaphors and similes in particular, must therefore involve property attribution rather than property matching. For literal comparisons, properties of a predicate that are informative about and relevant to a subject are attributed to that subject. For metaphors and similes, properties of the vehicle are attributed to the topic. This observation is consistent with Glucksberg and Keysar's (1990) portrayal of metaphoric statements as class-inclusion assertions. On this view, *Deanna's job can be described as being "like" a prison because it "is" a "prison,"*

in the sense that it belongs to a category of unpleasant, confining situations that jails exemplify. Abstract categories of this sort provide the properties that are attributed to the topic and serve as the basis for both the class-inclusion and the comparison assertion.

APPENDIX A

Materials for Experiment 1

Warm-up items and fillers: The brain is like a computer; The Middle East is a powder keg; Windows are like social classes; The AK-47 is an assault rifle; Some tranquilizers are like lectures; Tangerines are like oranges; All vegetables are eggplants; My collie was a sheepdog; My minister was a Lutheran; All fruits are apples; His therapy was psychoanalysis; Our president is a Republican; Some oranges are birds; Travelling is my hobby; Premeditated murder is a crime; Her feet are like paperclips; Russia is a communist country; Football is a professional sport; Paper is lemon juice; Sharp knives are weapons; Her art is a business; All birds are sparrows; His investment is real estate; Some wood is steel; His job is like a watermelon; Churches are Christian religious buildings; Books are carpets; All animals are horses; Some pianos are pencils; Late dinners are like guitars; Computers are crocodiles.

Metaphors (similes): My marriage was an icebox; His lawyers were sharks; His heart was ice; His landlord was a vampire; My surgeon was a butcher; Inner-city schools are zoos; Salesmen are high-powered bulldozers; Billboards are like large warts; Chronic crime is a disease; Her words were daggers; Her education was a ladder; His anger is a volcano; Our memory is a sponge; Reagan's smile was a magnet; Love was the couple's voyage; Cigarettes are her pacifiers; Their trust is glue; Her beauty was a passport; Edison's ideas were diamonds; Faith was Ghandi's fortress; A person's best friend is an anchor; Professional soldiers are pawns; His obligations are shackles; Jealousy is an uncontrollable infection.

Literat comparisons: His cat is like a

dog; An apartment is like a condominium; Mosques are like African Churches; A Ph.D. is like his M.D; Their president is like a Prime Minister; Oriental pearls are like opals; Cheap scotch is like gin; Southern Mexico is like Spain; Stock brokers are like smart bankers; Envy is like excessive greed; This priest is like a congregation's minister; Linen is like expensive cotton.

APPENDIX B

Materials for Experiment 2

Low constraint topics: Life; People; Family; Earth; Country; Apartment; Vacation; Crime; Dancer; School; Animal; Brother.

High constraint topics: Plastic Surgeon; Computer Program; Beggar; Memory; Jealousy; Mind; Idea; Smile; Temper; Friendship; Experience; Beauty.

Ambiguous Vehicles: Garden; Iceberg; Desert; Journey; Virus; Jungle; River; Ocean; Tree; Television; Puppy; Filing Cabinet.

Unambiguous Vehicles: Prison; Crutch; Doorway; Earthquake; Dagger; Timebomb; Fortress; Sleeping Pill; Weapon; Textbook; Racehorse; Goldmine.

Low-constraint topic metaphors: His life is a soap opera; Some people are puzzles; Her family is an anchor; The Earth is a beehive; Some countries are pressure cookers; Some apartments are palaces; A vacation is medicine; Crime is a disease; Dancers are butterflies; Some urban schools are zoos; Some animals are bodyguards; His brother is a mouse.

High-constraint topic metaphors: Some plastic surgeons are butchers; Some computer programs are labyrinths; Beggars are parasites; Some memories are snapshots; Jealousy is an infection; The mind is a landscape; Some ideas are diamonds; A smile is a magnet; Some tempers are volcanoes; Some friendships are glue; Experience is a fountain; Beauty is a passport.

Ambiguous vehicle metaphors: Some colleges are gardens; Some offices are icebergs; Some relationships are deserts; Love is a journey; Rumors are viruses; Shopping malls are jungles; Some dreams are rivers; Children are

puppies; Wisdom is an ocean; Old men are trees; Young girls are televisions; Some homes are filing cabinets.

Unambiguous vehicle metaphors: Some jobs are prisons; Alcohol is a crutch; An education is a doorway; Some divorces are earthquakes; Angry words are daggers; Cigarettes are timebombs; Faith is a fortress; Some lectures are sleeping pills; Sarcasm is a weapon; Some professors are textbooks; Some teenagers are racehorses; Some stocks are goldmines.

REFERENCES

- ARISTOTLE (1965). *Poetics*. Oxford: Oxford Univ. Press.
- BLACK, M. (1962). *Models and metaphors*. Ithaca, NY: Cornell Univ. Press.
- BLACK, M. (1979). More about metaphor. In A. Ortony (Ed.), *Metaphor and thought*. (pp. 19–43). Cambridge: Cambridge Univ. Press.
- BRACHMAN, J. R. (1983). What IS-A is and isn't: An analysis of taxonomic links in semantic networks. *Computer*, **16**, 30–36.
- BROWN, R. (1958). *Words and things*. New York: Free Press.
- CLARK, H. H., & HAVILAND, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension*. Norwood, NJ: Ablex.
- GENTNER, D. (1980). *Studies of metaphor and complex analogies*. Paper presented at the annual meeting of the American Psychological Association. Montreal, Canada.
- GENTNER, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, **7**, 155–170.
- GIBBS, R. J. (1984). Literal meaning and psychological theory. *Cognitive Science*, **8**, 275–304.
- GLUCKSBERG, S., GILDEA, P., & BOOKIN, H. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, **21**, 85–98.
- GLUCKSBERG, S., & KEYSAR, B. (1990). Understanding metaphoric comparisons: Beyond similarity. *Psychological Review*, **97**, 3–18.
- GOODMAN, N. (1972). *Problems and projects*. New York: Bobbs-Merrill.
- HARRIS, R. (1976). Comprehension of metaphor: A test of a two-stage processing model. *Bulletin of the Psychonomic Society*, **8**, 321–324.
- HARWOOD, D. L., & VERBRUGGE, R. R. (1977). Metaphor and the asymmetry of similarity. Paper presented at the convention of the American Psychological Association.
- KATZ, A. N. (1982). Metaphoric relationships: The role

- of feature saliency. *Journal of Psycholinguistic Research*, **11**, 283–296.
- KEYSAR, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language*, **28**, 375–385.
- MEDIN, D. L., GOLDSTONE, R. L., & GENTNER, D. (1993). Respects for similarity. *Psychological Review*, **100**, 254–278.
- MILLER, G. A. (1979). Images and models; Similes and metaphors. In A. Ortony (Ed.), *Metaphor and Thought*. Cambridge: Cambridge University Press, pp. 202–250.
- MURPHY, G. L. (1988). Comprehending complex concepts. *Cognitive Science*, **12**, 529–562.
- MURPHY, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, **29**, 259–288.
- 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**, 465–478.
- ORTONY, A., VONDRUSKA, R. J., FOSS, M. A., & JONES, L. E. (1985). Saliency, similes, and the asymmetry of similarity. *Journal of Memory and Language*, **24**, 569–594.
- TRAGER, G. L. (1936–1939). “Cottonwood-Tree”, a southwestern linguistic trait. *International Journal of American Linguistics*, **9**, 117–118.
- TVERSKY, A. (1977). Features of similarity. *Psychological Review*, **85**, 327–352.
- VERBRUGGE, R., & MCCARRELL, N. (1977). Metaphoric comprehension: Studies in reminding and resembling. *Cognitive Psychology*, **9**, 494–533.
- WECHSLER, D. (1958). *The measurement and appraisal of adult intelligence (4th ed.)*. Baltimore: Williams and Wilkins.
- 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: LEA.

(Received February 16, 1995)

(Revision received January 23, 1996)