Birds of a Feather Flock Together: Similarity Judgments with Semantically Rich Stimuli

MIRIAM BASSOK
University of Chicago

AND

DOUGLAS L. MEDIN
Northwestern University

The structural alignment approach to similarity posits a principled distinction between object attributes and relations between objects. We examined whether this assumption holds for metaphorical combinations of (interrelated) objects. Subjects judged similarity between simple statements in which the nouns (denoting attributes) and verbs (denoting relations) were semantically interdependent. We found that semantic dependencies affected similarity judgments both by inducing inferences about the abstract combined meaning of the statements and by changing the process by which subjects arrived at their judgments. When the paired statements had matching verbs (e.g., "The carpenter fixed the chair" and "The electrician fixed the radio"), subjects compared the combined meanings of the statements (e.g., "Similar because both are professionals doing their job"). These results are consistent with the logic of structural alignment. However, when the paired statements had matching nouns (e.g., "The carpenter fixed the chair" and "The carpenter sat on the chair"), very often subjects integrated the combined meanings of the statements (e.g., "Similar because he sat on the chair to see whether he fixed it well"). These results defy every existing account of similarity. We discuss the prevalence and systematicity of such processing replacements and the need for incorporating them into similarity-based accounts of cognition.

© 1997 Academic Press

Similarity serves as a central explanatory construct in theoretical accounts of many cognitive processes (e.g., categorization, transfer of learning, probabilistic reasoning). However, it remains unclear how people determine similarity between objects, situations, or events. The prevalent definition of similarity between two entities is the degree of overlap between their respective properties. In particular, following Tversky’s (1977) contrast model, it is commonly agreed that similarity is computed by some monotonic function that assigns positive weights to properties that are common to the compared entities (i.e., matches) and negative weights to properties that are unique to one of the compared entities (i.e., mismatches). The gist of this definition is consistent with people’s intuitive understanding that common properties increase similarity (e.g., having team members wear uniforms) whereas distinctive properties decrease similarity (e.g., adding distinct color patches to white outfits of tennis players). Yet, in itself, this definition is insufficient to capture the pattern of similarity judgments. In fact, Medin, Goldstone, and Gentes (1993) review a large body of evidence showing that similar-

This work was supported by Grant BNS 95-11757 from the National Science Foundation and by a grant from the University of Chicago School Mathematics Project. Portions of this research were presented at the Seventeenth Annual Conference of the Cognitive Science Society, July 1995, Pittsburgh, Pa. We thank Larry Barsalou, Art Mackman, Tom Ward, and especially Ed Wiśniewski for their insightful comments on an earlier version of the manuscript. Address correspondence and reprint requests to Miriam Bassok, Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637. E-mail: mbassok@uchicago.edu.

0749-596X/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.
by judgments are highly sensitive to a variety of task characteristics (e.g., context, task in-
suctions) and therefore, judging by this
definition, they may appear to be "disturbin-
gly flexible" (p. 258).

The high variability in people's similarity
judgments suggests that inferential mecha-
nisms external to similarity (e.g., categoriza-
tion) determine the relevant "respects" for
similarity, i.e., the relevant properties that
are subsequently matched and weighted in the
computation of similarity (e.g., Goodman,
1972; Murphy & Medin, 1985; Tversky,
1977). Nevertheless, Medin et al. (1993) ar-
gued that the observed variability in similarity
judgments disguises systematic regularities in
the process of comparison by which people
actively construe similarity. In particular, they
proposed that similarity judgments are medi-
ated by a process of structural alignment that
imposes systematic constraints on the selec-
tion and weighting of the relevant properties.

As we elaborate shortly, the alignment frame-
work draws a principled distinction between
matches and mismatches in two types of prop-
erties: object attributes and relations between
objects. At present, however, evidence for the
existence of processing regularities associated
with this distinction is limited to stimuli in
which these two types of properties are separa-
table and independent (e.g., a star is above a
circle). In this paper we report results from
three experiments that examined whether the
processing principles of the alignment frame-
work can be extended to stimuli consisting of
semantically constrained combinations of interrelated objects, i.e., simple noun–verb–
noun statements such as "The carpenter fixed
the chair," "The electrician fixed the radio,"
and "The carpenter sat on the chair."

We began this work with the hypothesis
that semantic dependencies between nouns
denoting object attributes) and verbs (denot-
ing relations between objects) will affect simi-
larity judgments by leading people to include
in their representations inferences that refer to
the abstract combined meaning of the state-
ments (e.g., a professional performing a job-
related or a job-unrelated activity). Indeed, by
 coaxing such semantically induced inferences
as higher order relations, we successfully pre-
dicted the pattern of similarity ratings in a way
that is consistent with the logic of structural
alignment. However, to our great surprise, the
justifications that accompanied people's rat-
ings strongly suggest that comparison is not
the only process by which people determine
similarity. Specifically, we found that people
do not always compute the degree of overlap
between the properties of the paired stimuli.
Rather, contrary to any theoretical account of
similarity, for certain stimuli they systemati-
cally integrate the paired stimuli into a com-
mon thematic scenario (e.g., "Quite similar
because the carpenter sat on the chair to see
whether he fixed it well"). That is, we found
that when people are asked to judge similarity,
they may systematically replace the process
of comparison with a process to which we
typically refer as thematic association. Before
developing this point we first set the context
for the present studies in terms of an alignment
approach to similarity comparisons.

The Structural-Alignment Hypothesis: Alignment of Attributes and Relations

The structural-alignment hypothesis was
originally proposed by Gentner (1983) to ac-
count for the process of analogical mapping.
The gist of the structural-alignment hypothesis
is that, in order to determine whether two enti-
ties are analogous, people attempt to bring
into correspondence the representations of the
compared stimuli in a way that will result in
the most globally coherent alignment. The
main assumption of this theoretical framework
is that analogical mapping is mediated by
matches and mismatches in two qualitatively
different types of properties: 1. Attributes,
one-place predicates that take objects as argu-
ments (e.g., RED [apple]), and 2. Relations,
predicates that take two or more attributes,
objects, or other relations as arguments (e.g.,
SAME COLOR [apple, book]). This assump-
tion has been incorporated into various com-
putational models of analogical mapping (e.g.,

Gentner’s (1983) “systematicity principle” states that relational matches constrain the alignment of attributional matches, and that higher order relations (i.e., relations that take other relations as arguments) constrain the alignment of lower order relations (i.e., relations that take objects as arguments). Following alignment, relational matches receive higher weights than attributional matches. According to Gentner’s (1983) structure-mapping theory, the distinction between and the dominance of relational matches over attributional matches, is important because what matters to analogical reasoning is finding a common relational structure in the compared base and target stimuli, whereas matches and mismatches in attributes of the specific objects can be neglected. For example, when one suggests that “an atom is like the solar system,” the intended meaning involves matches in relations between objects such as X revolves around Y or X attracts Y, rather than matches in attributes of the specific object such as X is hot or Y is yellow.

More recently investigators have entered into the idea that similarity is like analogy with respect to alignment and differential weighting of attributes and relations. Although similarity does not demand relational commonalities, it demands alignment. Hence, by assuming that people always align stimuli in the same way, i.e., always distinguish between attributional and relational matches, the structural-alignment hypothesis could serve as a general mechanism that accounts for both similarity judgments and analogical mapping. This logic motivated recent attempts to adapt the processing assumptions of the structural-alignment hypothesis to the more general class of similarity judgments (e.g., Goldstone, 1994; Goldstone & Medin, 1994; Goldstone, Medin, & Gentner, 1991; Markman & Gentner, 1993, 1996; Medin et al., 1990, 1993).

The work of Goldstone et al. (1991) provides a good example of empirical evidence that supports the psychological validity of the distinction between relations and attributes in similarity judgments. Using stimuli consisting of interrelated geometric objects such as (1) “a triangle is above a square,” (2) “a star is above a circle,” or (3) “a triangle is next to a circle,” they asked people to choose which of two target alternatives (e.g., 2 or 3) was more similar to a standard base (e.g., 1). The stimuli were constructed such that one target had an extra relational match and the other an extra attributional match with the base (e.g., “above” and “triangle” in targets 2 and 3, respectively). In various conditions further relational or attributional matches were added to both targets. The proportion of relational choices depended on the overall context—the target with an extra relational match was more likely to be selected when both targets shared numerous relations with the base than in contexts where both targets shared numerous attributes with the base. Goldstone et al. (1991) developed the MAX model to account for such nonlinear weighing of relational and attributional matches in similarity judgments. The MAX model postulates that people process these two types of matches in separate “pools” and assign higher weights to matches in the bigger pool.

In another study (Medin et al., 1990), the same authors observed an interesting phenomenon that further underlines the distinction between relational and attributional matches. Using stimuli similar to those used by Goldstone et al. (1991), they found that when the task was to judge similarity, subjects were more likely to choose targets with relational rather than with attributional matches as being more similar to the base. However, when the task was to judge difference, this preference was less pronounced or even reversed. As a result, in an extreme case, subjects presented with a base such as OXO chose a relational target (e.g., Δ*Δ, one relational match in symmetry) over an attributional target (e.g., Δ*O, one attributional match in circle) as both more similar to and more different from the base. Such nonmirroring judgments suggest
that people assign higher weights to relational matches in similarity than in difference judgments. At present, the reason for such differential weighting of relational and attributional matches in response to the type of judgment remains unclear (but see A. Markman, 1996). Nevertheless, results from this line of research strongly support the assumption of the structural-alignment hypothesis that there are systematic regularities in the process by which people arrive at similarity judgments, and that these regularities depend on the distinction between relational and attributional matches.

**Nonarbitrary Dependencies Between Relations and Attributes**

The stimuli used in the above studies were carefully chosen to enable a clear contrast between the relative impact of relational and attributional matches on similarity and difference judgments. In particular, the stimuli consisted of objects and relations that were separable and independent. For example, the symmetry of OXO and Δ
\*Δ
 (relations) is independent of whether the symmetric shapes in the two figures happen to be circles or triangles (attributes) or whether a star is between two triangles rather than vice versa (Δ
\*Δ
). However, people typically encounter entities in which objects and their attributes are interrelated in a nonarbitrary way. That is, in most cases it matters which objects serve as arguments in a particular relation. For example, carpenters are known to fix chairs rather than radios as part of their profession; cutting grass is a different type of cutting than cutting hair; and mockingbirds are believed to sing more nicely than crows. There is substantial evidence showing that people spontaneously draw inferences based on their knowledge about the way in which various objects tend to be interrelated (e.g., Anderson & Ortony, 1975; Gentner & France, 1988; Ortony, 1979; Strohner & Nelson, 1974). Given such spontaneous inferences, it is quite possible that the assumption of the MAX model, that people process attributional and relational matches in separate pools, does not apply to base and target stimuli in which these two types of properties are semantically interdependent.

The impact of nonarbitrary dependencies between attributes and relations has been recently examined in the context of analogical problem solving (see Bassok, 1996, for a review), where it is commonly accepted that relational matches are important because they are solution relevant whereas attributional matches are superficial to problem solutions. These studies have shown that people spontaneously use semantic dependencies between objects and relations to interpret the structure of analogous problems, and that matches between such “interpreted representations” of the problems can override explicit relational and attributional matches.

For example, subjects in the Bassok and Overholt (1995) study spontaneously interpreted the meaning of the predicate constant change(x) using their knowledge about the manner (continuous or discrete) in which x is changing, and their interpretations affected the scope of analogical transfer. Specifically, college students who learned to solve physics problems involving constant change in the speed of moving objects spontaneously interpreted the situations described in the problems as cases of continuous change. As a result, 71% of subjects transferred the physics solution to analogous nonphysics problems which they also interpreted as cases of continuous change (e.g., constant change in the rate of population growth in people/year), but only 27% of sub-
jects transferred the continuous physics solutions to problems that they interpreted as cases of discrete change (e.g., constant change in the rate of attendance at an annual fair in people/year). Such differential transfer from physics occurred even though the descriptions of the continuous and discrete target problems used identical relational terms (i.e., constant change) and object attributes (i.e., people, years). That is, transfer was mediated by matches in the combined interpreted representations of the problems rather than by the explicitly stated separate relational and attributional matches.

Bassok, Wu, and Oseh (1995) examined the impact of dependencies between relations and attributes on analogical mapping. In one experiment, subjects learned to solve a permutation problem in which a manager of a country club randomly assigned caddies to golfers. Then they were asked to solve a structurally isomorphic problem in which a manager either randomly assigned caddies to carts or carts to caddies. The structural-alignment hypothesis predicts that people should align caddies with caddies (i.e., identical objects) rather than carts with caddies (i.e., different objects) and therefore be more successful in solving the “caddies assigned to carts” than the “carts assigned to caddies” problem. However, contrary to this prediction, 94% of subjects solved correctly the “carts assigned to caddies” problem, but only 24% of subjects solved correctly the “caddies assigned to carts” problem. This pattern of results indicates that subjects spontaneously interpreted the assignment problems using their knowledge that golfers get caddies and caddies get carts. That is, they understood that the problems had the interpreted cart (receivers, givers) rather than the explicit assign (assignees, assigned) structures. As a result, they aligned givers with givers (carts with caddies) and receivers with receivers (caddies with golfers).

The studies by Bassok and her colleagues document that dependencies between relations and attributes can affect analogical transfer by inducing inferences that affect the representations constructed for the compared problems. However, it remains unclear whether such dependencies actually change the process by which people align the interpreted representations of the compared stimuli. In fact, it is possible that after interpreting the base and target problems, subjects proceeded by aligning the interpreted objects (e.g., receivers) and relations (e.g., get) just as they tend to do when aligning independent objects (e.g., triangle) and relations (e.g., above). Thus, at least in principle, these results are consistent with the processing assumptions of the structural-alignment hypothesis.

In this paper we present results of three experiments that examined whether and how dependencies between relations and attributes affect similarity judgments. To allow for inferential processes and nonindependent encoding of attributes and relations, the stimulus materials were triplets of simple statements in which the nouns (denoting objects) and verbs (denoting relations) were semantically interdependent. The triplets of statements were designed in a way that allowed us to distinguish effects of separate matches in the explicitly stated nouns and verbs from more abstract correspondences associated with the inferred combined meaning of the statements. We collected both similarity and difference judgments to assess effects of type of judgment on comparison and weighting. In so doing, we could examine whether the processing regularities that were found to result in nonmirroring similarity and difference judgments in the independent case (Medin et al., 1990) can be generalized to stimuli in which attributes and relations are interdependent.

Overview of Experiments

We constructed 12 quintuplets of simple noun–verb–noun statements, letting nouns stand for objects and their attributes and verbs for relations between the objects. Each quintuplet consisted of a base, an “attributitional target” that matched the nouns in the base, and three “relational targets” that matched the verb in the base. Using the 12 quintuplets
of statements, we constructed three types of triplets, each consisting of a base, its attributional target, and one of its three relational targets. For each triplet, subjects had to choose whether the relational or the attributional target was more similar to (different from) the base or to rate the similarity (difference) between the base and its relational and attributional targets. Below is one quintuplet of statements that exemplifies the design of our stimuli:

**Base:** The carpenter fixed the chair

**Attributional target (AA):**

The carpenter set on the chair

**Relational targets:**

Relation = is-a Match Relation* (RR*)

The electrician fixed the radio

Relation (R):

The plumber fixed the radio

Relation = Attribute (RA):

The carpenter fixed the radio

The target statements were constructed such that they matched the base in their nouns, verbs, or both (set in boldface). Hereof we refer to these attributional (A) and relational (R) matches in nouns and verbs as separate matches. In addition to such separate matches, the targets either matched or did not match the base in their inferred combined meaning. We assume that matches in the combined meaning of the statements can be rated as matches in abstract higher order notions (denoted by R*) that specify how separate attributes and relations are interrelated.

The combined meaning of the RR* targets always matched the combined meaning of the base (e.g., a professional performing a job-related activity), whereas the combined meaning of the R or RA targets did not (e.g., neither plumbers nor carpenters fix radios as part of their profession). As a result, although the three relational targets had the same separate relational match with the base (e.g., R = fix), RR* formed good analogies to the base whereas R and RA formed poor analogies. The difference between the poor analogies: R and RA targets was that RA had an additional attributional match with the base (e.g., A = carpenter). The attributional target (AA) had a relational mismatch with the base (e.g., fix + sit) and therefore a combined mismatch.

Each of the three relational targets from a given quintuplet was always paired with the same attributional target. Hence, by comparing subjects' choices or ratings for the three types of relational targets, we could examine the relative impact of separate and combined matches on similarity and difference judgments. Our predictions reflect an attempt to extend the logic of the structural-alignment hypothesis to the interpreted representations of stimuli in which attributes and relations are semantically interdependent. Specifically, by assuming that matches in the abstracted combined meanings of the statements can be treated as matches in higher order relations, the logic of Gentner's (1983) systematicity principle dictates that combined matches should be weighted more highly than separate relational or attributional matches. Accordingly, we predicted that RR* targets will be rated as more similar to the base than R targets even though both targets are equated in their explicit separate matches to the base (R). Moreover, we predicted that RR* targets will be rated as more similar to the base than RA targets because the inferred higher order relation (R*) should have more impact on similarity judgments than an additional separate attributional match (A). To the extent this similarity is a monotonic function of matching properties, the overall pattern of similarity ratings for the three relational targets should be RR*> RA > R. Contrast predictions are apparent if, as in the independent case, separate matches are the only factor influencing judgments. In that event, RA (two separate matches) should be more similar to the base than either RR* or R, which should not differ from each other (one separate match).

As mentioned earlier, we collected both
similarity and difference judgments to examine whether and how separate and combined matches interact with the type of judgment. The pattern of nonmirror effects observed by Medin et al. (1996) suggests that relational matches are weighted more highly in similarity than in difference judgments and attributional matches are weighted more highly in difference than in similarity judgments. If this processing regularity also holds for the interpreted representations of the base and target statements, then the relational targets should be chosen more often as more similar to the base than as less different from the base. Moreover, by assuming that the RR* relational target has an inferred higher-order relational match available to be differentially weighted in similarity and difference judgments, we anticipated that the nonmirror effects will be stronger for the RR* than for the R and RA targets. It is less clear what one should predict about the relative magnitude of nonmirror effects in triplets with the nonanalogous relational targets. In Goldstone’s (1994) SIAM model, the weight of mistratches increases with the strength of alignment. From this perspective one might expect stronger nonmirrors for RA than for R-only targets (though SIAM has not been elaborated to deal with similarity versus difference judgments).

The three experiments used the same stimuli and the same experimental design but differed in the dependent measures of similarity and difference. In Experiment 1 subjects were asked to choose whether a relational or an attributional target was more similar to (or different from) a given base; in Experiment 2 they were asked to rate the degree to which each target was similar to (or different from) a given base; and in Experiment 3 they were asked to both rate similarity or difference and explain in writing how they arrived at their ratings.

**Experiment 1**

**Method.** Subjects. Subjects were 84 undergraduates from Northwestern University. They were tested individually or in small groups and paid for their participation. **Materials.** As described earlier, a total of 60 noun-verb-noun statements were constructed for this study: 12 base statements, each with its corresponding attributional target (AA) and three relational targets (RR*, RA, and R). Eight of the 12 quintuplets of statements described a functional relation between an actor and an object (action statements, e.g., "An engineer designed a car"), and 4 compared an attribute of two entities (comparison statements, e.g., "A cathedral is higher than a church"). The 12 sets of base and target statements appear in the Appendix.

Using the 12 quintuplets of statements we constructed three types of triplets, each consisting of a base, its attributional target, and one of its three relational targets. Each triplet was typed on a separate page, with the base centered above the two targets. The left and right position of the relational and attributional targets were randomized. We constructed 12-page booklets by randomly selecting a triplet type for each of the 12 base statements and then collating the selected pages in randomized order. Half of the booklets had a cover page asking subjects to choose which of the two bottom statements was more similar to the top statement, and the other half had a cover page asking them to choose which of the bottom statements was more different from the top statement.

**Procedure.** Subjects were randomly assigned to the similarity (N = 42) and difference (N = 42) conditions and made their choices at their own pace. The task lasted about 10 min.

**Results.** For each subject we computed three scores, one for each triplet type. The scores were the relative proportions of items within a triplet type in which the relational targets were chosen as more similar to the base for similarity judgments or as less different from the base (i.e., not chosen as more different) in difference judgments. Figure 1 presents the average
percentages of relational choices for each of the three relational targets in the similarity and difference conditions. The results are collapsed across the action and comparison statements following a preliminary analysis that did not show any significant effects due to statement type.

Overall, collapsing across the similarity and difference conditions, the three triplet types differed in the frequency with which subjects chose the relational target as more similar to or less different from the base (F(2,82) = 6.84, p < .002). As predicted, relational choices were most frequent for targets with an inferred combined match (RR* = 63%), less frequent for targets with an additional explicit attributional match (RA = 56%), and least frequent for targets with a single explicit relational match (R = 48%), that is, RR* > RA > R. Planned comparisons showed that relational choices for each of the triplet types differed reliably from the other two types, and this difference held true for most of the items. Specifically, even though RR* and R were equated in their separate matches to the base, RR* targets (i.e., good analogies) were chosen more frequently than R targets (i.e., poor analogies) for 11 of the 12 items, with 1 item yielding equal choices. Moreover, the inferred combined match in good-analogy targets had more impact on similarity and difference judgments than an additional explicit attributional match in poor-analogy targets (RR* > RA) for 9 of the 12 items, and it was reversed only for 1 item.

The extent to which relational choices differ for similarity and difference judgments provides an index of nonmirroring effects. Unlike the independent case of Medin et al. (1990), the main effect of similarity versus difference judgment did not prove to be statistically reliable (F(1,82) = 2.29, p = .13). Moreover, although the overall pattern of relational choices (i.e., RR* > RA > R) appeared to be more prominent in the similarity than in the difference condition (72%, 61%, 48% vs 54%, 51%, 48%, respectively), the judgment by Triplet Type interaction was only marginally reliable (F(2,164) = 2.42, p = .09). Planned comparisons revealed that the nonmirroring effect of judgment type was significant only for the RR* target (72% more similar vs 54% less different), F(1,82) = 4.84, p < .04. Thus, we replicate the Medin et al. (1990) findings only in the condition where the combined interpretations of the base and target
statements would lead to an incorrect matching higher order relation (R^o).

Discussion

The results of Experiment 1 reveal a pattern of choice that cannot be predicted simply on the basis of explicit separate matches. Rather, as predicted, it appears that subjects constructed combined representations for the compared stimuli and relied on matches in the inferred combined meanings of the base and target statements to determine similarity. Although the inferred combined matches were implicit, subjects were more likely to choose good rather than poor-analogy targets when the two targets were equated in the number of explicit separate matches (RR^o > R), and even when the poor-analogy targets had an additional explicit separate match (RR^o > RA). For example, they were more likely to choose “The electrician fixed the radio” than “The carpenter fixed the radio” as more similar to the base “The carpenter fixed the chair,” even though a carpenter is clearly more similar to a carpenter than is an electrician. The stronger impact of inferred combined matches relative to explicit separate matches was also apparent in the pattern of choices in the similarity and difference conditions. Specifically, unlike the nonmirroring effects observed in the independent case (Medin et al., 1990), relational choices were reliably more frequent in the similarity than in the difference condition only in triplets with the good-analogy targets (RR^o).

The results of Experiment 1 are consistent with the findings of Bassok and her associates who have shown that dependencies between relations and attributes affect analogical transfer (Bassok & Osheth, 1995; Bassok et al., 1995). These findings underline the importance of interpretive processes that affect similarity judgments by affecting the psychologically relevant representations constructed for the base and target stimuli. At present, such interpretive effects remain outside the scope of the alignment models. As a result, similarity judgments might appear highly variable even when, in fact, there are systematic regularities in the process by which people construe similarity. Indeed, despite the significant departure from the pattern of judgments observed for combinations of separable and interdependent attributes and relations, it appears that the present pattern of results could be accommodated within the general framework of the structural-alignment hypothesis. Specifically, by treating combined matches as matches in higher order relations, as we have proposed in our initial analysis, the fact that matches in the inferred higher order relations override explicit separate matches is consistent with the logic implied by Gentner’s (1983) systematicity principle.

Attempts to extend the processing principles of the structural-alignment hypothesis to stimuli involving dependencies between attributes and relations may require further specification of some basic assumptions about processing. In particular, the fact that we observed nonmirroring effects only in triplets with the good-analogy targets (RR^o) suggests that people did not assign different weights to separate relational and attributional matches in similarity and difference judgments. Yet, it is possible that our measure of choice was not sensitive enough to capture nonmirroring weighting of separate relational and attributional matches. For example, even if people assigned higher weights to relational matches in similarity (R = 3) than in difference judgments (R = 4), such nonmirroring weighting could have been insufficient to show up in relational choices (e.g., given A = 3, both 5 > 3 and 4 > 3). Experiment 2 examined this possibility by decomposing the measure of choice into its two constituent ratings of the relational and the attributional targets.

Experiment 2

Experiment 2 replicated the design of Experiment 1, but instead of subjects being asked to choose between the relational and the attributional targets, they were asked to rate the extent to which each of the paired targets was similar to the base (or different from the base).
By soliciting direct judgments for the relational and attributional targets in each triplet, that is, by decomposing the choice task, we could examine more directly whether separate relational and attributional matches receive different weights in similarity and difference judgments. Furthermore, we could examine whether the pattern of choice observed in Experiment 1 reflected only differences in ratings of the relational targets (RR, RA, R) or some interaction between ratings of the relational and attributional targets. That is, even though the same attributional target was paired with each of the three relational targets, it is possible that such distinct pairings led to significant contrast effects (e.g., A\textsubscript{b} > A\textsubscript{a} > A\textsubscript{c}).

**Method**

Subjects. Subjects were 89 undergraduates from Northwestern University and from the University of Chicago. They participated in the experiment on a voluntary basis during their regular class periods.

Materials. The materials were identical with those used in Experiment 1 except that a 7-point rating scale appeared below each of the two target statements. The scale assessed either similarity or difference. Accordingly, the lowest rating on the scale (1) was labeled either “not at all similar” or “not at all different” and the highest rating (7) was labeled either “very similar” or “very different.” Also, the instructions on the cover page were modified to fit the rating task: Subjects were asked to rate how similar (different) each of the two bottom statements was to the top statement.

Procedure. Subjects were randomly assigned to the similarity (N = 42) and difference (N = 47) conditions and completed the booklets at their own pace. The task lasted about 10 min.

**Results**

For each subject we computed six rating scores by averaging the ratings of items for each of the three types of relational targets (RR, RA, R) and their corresponding attributional targets (A\textsubscript{b}, A\textsubscript{a}, A\textsubscript{c}). To enable comparisons of ratings obtained in the similarity and difference conditions we converted difference ratings into similarity ratings by subtracting them from 8. For example, the rating “not at all different” (1) was converted into “highly similar” (7). Table 1 presents the averages and standard deviations of these rating scores in the similarity and difference conditions. The results from the two types of items (i.e., action and comparison) and from the two schools were collapsed following a preliminary analysis that did not yield any significant effects due to these factors.

Overall, collapsing across the similarity and difference conditions, relational targets (M = 4.09) were rated as more similar to the base than attributional targets (M = 3.61; F(1,87) = 8.86, p < .004, but the magnitude of this effect differed in the three triplet types (F(2,174) = 14.99, p < .0001). As can be seen by comparing the top and bottom panels of Table 1, this interaction was solely due to differences in the ratings of the three relational targets (top panel of Table 1). Specifically, target RR (M = 4.52) was rated as more similar to the base than target RA (M = 4.24), which in turn was rated as more similar to the base than target R (M = 3.50). In each case, the dominance of the good analog target held true for 11 of the 12 items. That is, the pattern of similarity ratings for the three relational targets in Experiment 2 (i.e., RR > RA > R) replicates the pattern of relational choices in Experiment 1.

The ratings for the corresponding attributional targets (bottom panel of Table 1) did not differ across the three triplet types (F < 1). This lack of difference indicates that there were no significant contrast effects due to pairing of the same attributional targets with the three types of relational targets.

The pattern of nonmirroring similarity and difference ratings in Experiment 2 was qualitatively consistent with that observed for relational choices in Experiment 1. First, there was no difference between similarity and dif-
ference ratings of the attributional targets ($M = 3.61$ and $M = 3.61$, respectively, $F < 1$). Second, the relational targets were rated somewhat more highly in the similarity than in the difference condition ($M = 4.20$ and $M = 3.99$, respectively) although this trend was not significant ($F < 1$). Finally, the nonmirroring trend in the relational targets was most pronounced for the RR* good-analogy targets ($M = 4.72$ vs $M = 4.32$ in the similarity and difference conditions, respectively). However, unlike in Experiment 1, even this trend fell short of statistical reliability ($F(1, 87) = 2.10$, $p < .15$).

To promote comparison between Experiments 1 and 2, we computed the number of cases in which the relational targets received higher similarity and lower difference ratings than their corresponding attributional targets (ties were split evenly). This analysis assumes that whichever member of a pair received a higher similarity (lower difference) rating would have been chosen as more similar (less different), that is, gives an estimate of relational choices. The estimated frequencies of relational choices in Experiment 2 for the three triplet types in the similarity and difference conditions appear in Fig. 2. The pattern and the magnitude of the estimated relational choices in Experiment 2 was very similar to that in Experiment 1. Specifically, collapsing across the similarity and difference conditions, the estimated frequencies of relational choices for triplets with the RR*, RA, and R targets in Experiment 2 were 65%, 59%, and 49%, respectively. The corresponding frequencies of actual relational choices in Experiment 1 were 63%, 56%, and 48% (see Fig. 1). That is, it appears that the ratings in Experiment 2 would have resulted in a pattern of choice similar to that obtained in Experiment 1.

**Discussion**

The results of Experiment 2 show that the pattern of results obtained in Experiment 1 also holds when the forced-choice task is decomposed into direct ratings of the paired relational and attributional targets. We found that the ratings of the attributional targets remained stable across triplets and type of judgment. That is, differences between the three
triplet types were solely due to differences in ratings of the relational targets ($RR > RA > R$). Moreover, we did not replicate the non-mirroring effects observed in the independent case (Molin et al., 1990) even with the more sensitive measure of direct relational and attributional ratings. In particular, although there was a slight trend for the good-analogy targets ($RR^*$) to be rated more highly in the similarity than in the difference condition, there was no evidence for non-mirroring weighting of separate relational and attributional matches. Thus, as suggested in the discussion of Experiment 1, attempts to extend the processing principles of the structural-alignment hypothesis to the dependent case may require some adjustments to account for the lack of discrimination between separate relational and attributional matches in similarity and difference judgments.

**EXPERIMENT 3**

Experiment 3 was a replication of Experiment 2 except that subjects were asked to explain in writing their similarity and difference ratings. We were hoping that such explicit explanations would validate our assumption about inclusion of inferred higher order relations in the representations of the compared stimuli and shed further light on the nature of the process by which people arrive at similarity judgments.

**Method**

*Subjects.* Subjects were 80 undergraduates from Northwestern University and 90 from the University of Chicago. They were tested either individually or in small groups and were paid for their participation.

*Materials.* The materials were identical with those used in Experiment 2 except that the instruction page asked subjects to explain in writing how they arrived at their ratings. They were asked to explain all their ratings in the space provided below each rating scale.

*Procedure.* Subjects were randomly assigned to the similarity ($N = 39$) and difference ($N = 41$) conditions and completed the task at their own pace. The task lasted between 15 and 30 min.

*Results.*

*Similarity and difference ratings.* Similarity and difference ratings were scored and analyzed as in Experiment 2. Again, the results are collapsed across the two types of items.
(action and comparison) and across the two schools (following a preliminary analysis that did not reveal any significant effects due to these factors). Table 2 presents the averages and standard deviations of the similarity and the transformed difference ratings ($8 - rd = r$) for the three relational targets and their corresponding attributional targets in Experiment 3.

The pattern of ratings in Experiment 3 was fairly similar to that in Experiment 2 (see Table 1). In particular, collapsing across the similarity and difference conditions, there was a highly significant difference between the three triplet types ($F(2,156) = 22.79$, $p < .001$) and a highly significant interaction between triplet type and ratings of the relational vs. attributional targets ($F(2,156) = 35.75$, $p < .001$). As in Experiment 2, there was no difference between the three triplet types in the ratings of the attributional targets ($F < 1$), and the main source of difference between the three triplets was in the ratings of the three relational targets. Planned comparisons indicated that $R^*$ ($M = 4.63$) was rated as significantly more similar to the base than $R$ ($M = 3.92$), which in turn was rated as more similar to the base than $R^*$ ($M = 3.13$). That is, as in Experiment 2, the pattern of ratings for the three relational targets was $R^* > R > R^*$. Each dominance held true for 11 of the 12 subjects.

Comparison of Tables 1 and 2 suggests that there were some differences between Experiments 2 and 3 in the pattern of similarity vs. difference ratings (i.e., the pattern of nonmirroring effects). Specifically, in Experiment 3 there was not even a trend of nonmirroring effects for the good-analogy $R^*$ targets. Also, there was a marginally significant reversed trend for the poor-analogy $R$ targets to be rated more highly in the difference than in the similarity condition ($M = 3.56$ vs. $M = 2.89$, respectively, $F(1,78) = 2.78$, $p < .10$) and a nonsignificant trend for the attributional targets to be rated more highly in the similarity than in the difference condition ($M = 4.22$ vs. $M = 3.77$, respectively, $F(1,78) = 2.48$, $p = .12$). However, given that in both experiments all the nonmirroring trends related to judgment type were very weak and unreliable, we believe that any attempts to explain these trends will be highly speculative.

The main reliable difference between Ex-

<table>
<thead>
<tr>
<th>Relational targets</th>
<th>$N$</th>
<th>$M$</th>
<th>SD</th>
<th>$M$</th>
<th>SD</th>
<th>$M$</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity</td>
<td>39</td>
<td>4.58</td>
<td>1.18</td>
<td>4.83</td>
<td>1.44</td>
<td>2.89</td>
<td>1.24</td>
</tr>
<tr>
<td>Difference</td>
<td>41</td>
<td>4.68</td>
<td>1.23</td>
<td>3.82</td>
<td>1.39</td>
<td>3.56</td>
<td>1.26</td>
</tr>
<tr>
<td>Overall</td>
<td>80</td>
<td>4.63</td>
<td>1.21</td>
<td>3.92</td>
<td>1.41</td>
<td>3.13</td>
<td>1.26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attributional targets</th>
<th>$N$</th>
<th>$M$</th>
<th>SD</th>
<th>$M$</th>
<th>SD</th>
<th>$M$</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity</td>
<td>39</td>
<td>4.07</td>
<td>1.18</td>
<td>4.25</td>
<td>1.37</td>
<td>4.22</td>
<td>1.42</td>
</tr>
<tr>
<td>Difference</td>
<td>41</td>
<td>3.66</td>
<td>1.40</td>
<td>3.78</td>
<td>1.29</td>
<td>3.91</td>
<td>1.31</td>
</tr>
<tr>
<td>Overall</td>
<td>80</td>
<td>3.86</td>
<td>1.48</td>
<td>4.01</td>
<td>1.34</td>
<td>4.11</td>
<td>1.37</td>
</tr>
</tbody>
</table>

Note: $R$ denotes an explicit relational match (vert), $A$ denotes an explicit attributional match (novel), and $R^*$ denotes an inferred higher order relational match.

<table>
<thead>
<tr>
<th>$R^*$</th>
<th>$N$</th>
<th>$M$</th>
<th>SD</th>
<th>$R$</th>
<th>$M$</th>
<th>SD</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>$A_{R^*}$</th>
<th>$N$</th>
<th>$M$</th>
<th>SD</th>
<th>$A_R$</th>
<th>$N$</th>
<th>$M$</th>
<th>SD</th>
<th>$A_A$</th>
<th>$N$</th>
<th>$M$</th>
<th>SD</th>
</tr>
</thead>
</table>

TABLE 2
AVERAGE SIMILARITY AND CONVERSION DIFFERENCE RATINGS FOR THE THREE TYPES OF RELATIONAL TARGETS AND THEIR RESPECTING ATTRIBUTIONAL TARGETS IN EXPERIMENT 3
periments 2 and 3 was that in Experiment 2
the combined ratings of the relational targets
were significantly higher than those of the at-
tributional targets, but in Experiment 3 there
was no significant main effect of target type
($M = 3.89$ and $M = 3.99$ for relational and
attributitional targets, respectively, $F < 1$). As
can be seen by comparing Tables 1 and 2, this
difference reflects the fact that the attribu-
tional targets received somewhat higher rat-
ings in Experiment 3 than in Experiment 2 ($M
= 3.99$ vs $M = 3.61$, respectively), whereas
the relational targets received somewhat lower
ratings in Experiment 3 than in Experiment 2
($M = 3.89$ vs $M = 4.09$, respectively). That
is, asking subjects to explain their ratings in-
creased the magnitude of similarity ratings for
the attributional targets and decreased the
magnitude similarity ratings for the relational
targets.

The higher magnitude of attributional rat-
ings and lower magnitude of relational ratings
in Experiment 3 relative to Experiment 2 led
to a significant decrease in our measure of the
estimated proportion of relational choices (45
vs 58% in Experiments 3 vs 2, respectively).
Figure 3 depicts the estimated proportions of
relational choices in the similarity and differ-
ence conditions for the three relational targets
in Experiment 3. As can be seen by comparing
Fig. 3 with Fig. 2, the estimated frequency of
relational choices for each of the three rela-
tional targets (RR*, RA, and R) was lower in
Experiment 3 than in Experiment 2 (57%,
46%, and 31% vs 65%, 59%, and 49%, respec-
tively).

This difference between Experiments 2 and
3 suggests that when people judge similarity
between stimuli in which attributes and rela-
tions are serenitically interdependent, they
have a default preference to focus their atten-
tion on relational rather than on attributional
matches (Experiment 2). The request to ex-
plain their ratings seems to reduce this default
preference, possibly because subjects engage
in a more thorough processing (e.g., Smith &
Sloman, 1994) and therefore pay more atten-
tion to the nondefault attributional matches
and mismatches. This account is also consis-
tent with the finding that explanations nullified
the trend observed in Experiments 1 and 2 to
assign different weights to the inferred higher
order relations (RR*) in similarity and differ-
ence judgments. That is, nonmirroring
weighting of relational matches seems to re-
flect changes in the default focus of attention
during relatively superficial processing of the
compared stimuli.
To summarize, asking subjects for explanations affected the magnitude of their ratings. However, explanations did not affect the pattern of ratings, that is, $RR^* > RA > R$ and $AA_{SA} = AA_{SA} = AA_{SA}$. Hence, we believe that the explanations generated by the subjects can illuminate the nature of the process by which people construe similarity even when they do not explain their similarity and difference ratings.

Explanations of ratings. Two judges blind to the purpose of the experiment first parsed the explanation texts into separate "justifications" (i.e., reasons). They identified between 1 to 5 distinct justifications per target. This parsing resulted in a total of 1183 justifications for the relational targets ($M = 1.23$ per target) and 1223 justifications for the attributional targets ($M = 1.27$ per target). The judges then coded each justification into one of three categories: "syntactic," "separate," or "combined." In addition, they coded each justification as either supporting (+) or impeding (−) similarity. Each judge scored slightly more than half of the data, and interreliability was computed for the (random) overlapping subset of the materials (15%). The mean reliability was .84. The vast majority of the disagreements (80%) derived from explanations involving multiple justifications, where one judge might score an extra code as present that the other judge did not code. Table 3 presents the distribution of the syntactic, separate, and combined justifications generated for the relational and attributional targets in the three triplet types. Because the distribution of the justifications did not depend on type of judgment, the results reported in Table 3 are collapsed across the similarity and difference conditions.

Syntactic. Justifications were coded as syntactic when they involved syntactic labels (e.g., "different verbs"), reference to words (e.g., "only the words are the same"), or reference to the structure of the sentences (e.g., "both have the same structure"). As can be seen in Table 3, syntactic justifications were infrequent (3 - 5% per target type) and were distributed uniformly across the relational and attributional targets in all three triplet types. Hence, syntactic justifications cannot account for the nonuniform pattern of similarity ratings in the three triplet types.

Separate. Justifications were coded as separate when they referred to matching (+) and mismatching (−) aspects of the specific nouns and verbs in the base and target statements. This category included (a) references to the actual nouns and verbs (e.g., "both about a chair," "both fix") and (b) references to various implied properties of these nouns and verbs (e.g., "both race and piano competition involve victory," "cutting is actively doing something as opposed to passive thinking about something").

Separate justifications were very frequent (61 - 70% per target type). However, the distribution of separate matches (+) and mismatches (−) in the relational and attributional targets was virtually uniform. That is, with the possible exception of R triplets which induced a fewer number of separate matches, all targets induced about twice as many matches as mismatches. The difference in separate justifications raised for the relational and attributional targets in each triplet simply mirrored the structure of our stimuli: In the relational targets, separate matches referred to verbs and separate mismatches to nouns, whereas in the attributional targets, separate matches referred to nouns and separate mismatches to verbs. For the sake of clarity, the distinction between matches and mismatches in nouns and verbs is not reported in Table 3. Thus, despite their high occurrence, justifications that referred to separate matches and mismatches in nouns and verbs cannot explain the significant differences in similarity ratings (e.g., $RR^* > RA$, see Table 2).

Combined. The pattern of similarity ratings was adequately captured by justifications that referred to the inferred combined meaning of the base and target statements (27 - 35% per target type). In particular, the inferred combined matches and mismatches successfully captured the dominance of
TABLE 3

<table>
<thead>
<tr>
<th></th>
<th>Synoptic</th>
<th>Separate</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relational targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td>393</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RA</td>
<td>416</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>R</td>
<td>374</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Attributional targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>411</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A_0</td>
<td>407</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A_1</td>
<td>405</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: R denotes an explicit relational match (e.g.), A denotes an explicit attributional match (not), and R* denotes an inferred higher order relational match.

Good-analogy targets (RR*) over the poor-analogy targets (R, RA). Specifi- cally, sub- jects were more likely to mention combined matches (e.g., "Both involve professionals doing their jobs") when explaining their rat- ings for the RR* good-analogy targets (24%) than when explaining their ratings for the R (7%) and RA (8%) poor-analogy targets. Moreover, they were more likely to mention combined mismatches (e.g., "Not analogous because plumbers don’t fix radios as part of their job") when explaining their ratings for the R (26%) and RA (20%) poor-analogy tar- gets than when explaining their ratings for the RR* good-analogy targets (3%).

The relative frequency of combined justi- fications was similar for all target types. In particular, such justifications were as fre- quent in the explanations of the relational targets (30%) as in the explanations of the attributional targets (29%). However, the combined justifications generated for these two types of targets were qualitatively differ- ent. Specifically, to our great surprise, many of the combined justifications generated for the attributional targets did not compare the representations of the paired stimuli. Rather, these justifications integrated the representa- tions of the base and target statements into a common thematic scenario. Good scenarios (+) were raised as supporting similarity (e.g., "Quite similar because he sat on the chair to see whether he fixed it well"), and poor scenarios (-) were raised as impeding similar- ity (e.g., "To fix something doesn’t necessarily require use of it, a doctor fixes a heart but never uses it"). In order to further understand this surprising difference, we coded sepa- rately the combined justifications comprising the rightmost column of Table 3 as either "analogies" or "scenarios." Below are the coding criteria we used.

Analogies. Combined justifications were coded as analogies when they involved compari- son of the abstract combined meanings of the base and target statements. The most prev- alent type of justifications that were coded as analogies compared the salience appropriateness of the noun-verb-noun combinations in the base and target statements. For example: "A lawyer would never design a house" (unlike engineers who design cars), or "Most children do not enjoy jobs" (but they enjoy toys). Other combined justifications that were coded as analogies compared the informativeness, truth, or validity of the base and target state- ments. For example: "Does not say anything we don’t already know like the first one" (for
the base "Cottage cheese is cheaper than mozzarella cheese" and the target "Cottage cheese is cheaper than a house," or "The first sentence is a fact, whereas the second is an opinion." (for the base "Equations are more accurate than words") and the target "Equations are more difficult than words.") Some combined justifications in this category included explicit comments about the quality of the analogy (e.g., "The analogy is too far-fetched") or used a format typical to Analogies tests (e.g., "bacterich :: teen-lawn").

Scenarios: Combined justifications were coded as scenarios whenever they related the paired statements by a causal or a temporal relation. Examples of causal scenarios are: "A teacher may have listened to the lecture to prepare" (for the base "The teacher prepared a lecture") and the target "The teacher listened to the lecture."); "Something the child might do if he/she enjoyed the toy, out of selfishness" (for the base "The child enjoyed the toy") and the target "The child hid the toy."); or "A logical step: Since equations are more accurate they are more difficult to use" (for the base "Equations are more accurate than words") and the target "Equations are more difficult than words."). Examples of temporal scenarios are: "Very similar since he is cow with the product he has just flown"; "A child might hide the toy while playing with it"; "The barber would probably think about the hair before he cut it"; or "Examining a case is something the lawyer would do while talking it."

Again, two judges blind to the purpose of the experiment coded the data, and reliability was computed over a subset of the items (15%). Reliability was reasonably high for both the scenario category (88) and for the analogy category (85). As in the previous coding (Table 3), the major disagreements arose on justifications eliciting multiple codes. i.e., where one judge identified and coded more justifications than the other. Table 4 presents the relative proportions of analogies and scenarios generated for the relational and attributional targets in the three triplet types,

<table>
<thead>
<tr>
<th>n</th>
<th>Analogies</th>
<th>Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>106</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>117</td>
<td>87</td>
<td>13</td>
</tr>
<tr>
<td>132</td>
<td>97</td>
<td>3</td>
</tr>
</tbody>
</table>

Note: R denotes an explicit relational match (verb). A denotes an explicit attributional match (noun), and B+ denotes an inferred higher order relational match.

that is, it presents a breakdown of the combined justifications in the rightmost column of Table 3. Because the distribution of analogies and scenarios did not depend on type of judgment, the results reported in Table 4 are collapsed across the similarity and difference conditions.

As can be seen by comparing the top and bottom panels of Table 4, analogies were virtually the only type of combined justifications generated for each of the three relational targets, whereas more than half of the combined justifications generated for the attributional targets were scenarios. Overall, 94% of the 355 combined justifications generated for the relational targets were analogies and only 6% were scenarios. By contrast, only 30% of the 355 combined justifications generated for the corresponding attributional targets were analogies and 61% were scenarios. Importantly, scenario justifications were not unique to a handful of subjects. Rather, 73% of the 80 subjects in Experiment 3 generated at least one scenario justication. Moreover, scenario justifications were distributed uniformly among the 12 attributional items.

The difference in the distribution of analogies and scenarios for the relational and attributional targets was extremely systematic, suggesting that these two types of stimuli in-
ducis (or highlighted) two qualitatively different processes by which subjects arrived at their judgments. When the base and target
stacments shared a relational match (i.e., common verb), similarity and difference judgments were almost exclusively mediated by a process of structural alignment (analogy justifications). By contrast, when the base and target statements shared only attributonal matches (i.e., common nouns), similarity judgments were frequently mediated by a process of thematic integration (scenario justifications). Note that, consistent with this distinction, only the relational targets that matched the base both in their verb and in their noun (RA) led to a significant proportion of scenario justifications (13%).

As we have mentioned earlier, the distribution of combined matches (+) and mismatches (−) in the three types of relational targets (i.e., the distribution of analogy justifications) was consistent with the magnitude of similarity ratings of these targets (i.e., higher ratings for RR than for R and RA targets). This correspondence supports our hypothesis that matches in the inferred higher order relations affect similarity ratings. An interesting question is whether a similar correspondence exists between supporting (+) and impeding (−) scenarios and the ratings of the attributonal targets that were accompanied by such thematic justifications. Of course, this sort of post hoc analysis is subject to selection effects and is strictly correlational. Given this caveat, however, the data clearly suggest that thematic integration affects the magnitude of similarity ratings. Specifically, we found that similarity ratings were substantially higher for attributonal targets that were accompanied by supporting scenarios (M = 5.0) than for attributonal targets that were accompanied by impeding scenarios (M = 2.3). Attributional targets that were not accompanied by scenario justifications received intermediate ratings (M = 4.0). Thus, it appears that both types of combined justifications (analogies and scenarios) are valid indicators of the two qualitatively distinct processes (comparison and thematic integration) by which our subjects arrived at their similarity ratings.

Discussion

Explanations did not change the pattern of ratings relative to that observed in Experiment 2 (RR > RA > R) but increased the magnitude of attributional ratings and decreased the magnitude of relational ratings. As we have conjectured earlier, these changes in the magnitude of attributional and relational ratings suggest that the request to provide explanations led to a more thorough processing and therefore reduced effects caused by selective focus of attention on relational matches during more superficial processing of the stimuli. This account is also consistent with the finding that when asked to explain their ratings, even the inferred combined relational matches in the good-analogy targets (RR) were not rated more highly in the similarity than in the difference condition. Thus, it appears that monitoring similarity and difference judgments reflect changes in the default focus of attention on relational matches during superficial processing.

The written explanations provide direct evidence that people are sensitive to nonarbitrary semantic dependencies between attributes and relations and include the inferred combined meaning of the statements in their representations of the base and target stimuli. For example, they include inferences about role appropriateness of the specific objects as slot fillers of the relational terms (e.g., “Carpenters do not fix radios as part of their profession”). Also, consistent with our claim, justifications referring to the combined meaning of the statements were compatible with the different magnitudes of similarity and difference ratings. These results underline the importance of inferential processes that affect similarity judgments by allowing the representations constructed for the base and the target stimuli. Most strikingly, however, analysis of the combined justifications revealed that when the paired stimuli shared purely attributonal matches (i.e., the attributonal targets), people
did not necessarily form structural alignments. Rather, they often constructed thematic scenarios which affected their ratings (e.g., “First he broke [the chair] by sitting on it, or tested it after he fixed it.”). Given that thematic scenarios do not support similarity based on the relative proportion of matching and mismatching properties, our results indicate that some similarity judgments cannot be accommodated by any existing account of similarity.

**General Discussion**

We began this work with the goal of extending the structural-alignment approach to similarity to stimuli in which attributes and relations are semantically interdependent. Using simple noun-verb-noun statements as materials, we found that semantic dependencies between attributes and relations affected similarity and difference judgments both by inducing inferences that change the representations of the paired stimuli and by changing the process by which people construct similarity. Some of the processing differences between the dependent and the independent case were consistent with the logic that motivates the structural-alignment approach to similarity. In particular, the pattern of similarity and difference ratings observed for the relational targets and the explanations that accompanied these ratings were consistent with Gentner’s (1983) systematicity principle: Matches in the inferred higher order relations had more impact on similarity and difference judgments than matches in separate relations and attributes (i.e., RR > RA > R). The only reliable evidence for nonmirroring similarity and difference ratings came from target statements that matched the base in their inferred higher order relations (RR*). This change in processing from the independent case (i.e., lack of nonmirroring effects for the poor analogy targets) could be incorporated into some modified version of the structural-alignment hypothesis.

The pattern of similarity judgments observed for the relational targets was highly systematic and very robust—it held true in all three experiments despite changes in measures of similarity (i.e., forced choice, ratings, and explanations). This pattern supports the conjecture put forward by Medin et al. (1993) that variability in similarity judgments might be only apparent because it might reflect systematic regularities in the process by which people construe similarity (e.g., structural alignment). At the same time, however, our results also show that the high variability in people’s similarity judgments cannot be readily resolved even at the processing level—even when one takes into account changes to the representations constructed by people for the paired stimuli. In particular, analysis of the justifications generated by subjects in response to a request to explain their ratings (Experiment 3) revealed that similarity judgments were mediated by at least two qualitatively different types of processes: (a) a process that involves comparison of properties comprising the representations of the base and target stimuli and (b) a process that involves integration of the paired stimuli into a common thematic scenario.

The comparison process mediated similarity judgments for most of the relational targets, whereas thematic integration mediated similarity judgments for the majority of the attributional targets. Thus, it is not the case that similarity judgments vary only because other inferential mechanisms determine the relevant respects for similarity and change the weights assigned to matching and mismatching properties (e.g., inferences about combined matches derived from people’s semantic knowledge). Rather, our results indicate that similarity judgments can also vary because inferential processes external to similarity (i.e., thematic integration) actually replace the comparison process. Below we elaborate on this point and place our surprising findings in a broader context.

**Attractibility and Processing Replacements**

The reasoning that our subjects displayed in their thematic scenarios is strikingly similar to the “immature” reasoning of young children (see e.g., E. Markman, 1989, for a re-
view) or to that of illiterate adults from Uzbekistan before they had been educated by Soviet teachers (Luria, 1976). For example, one of the illiterate peasants interviewed by Luria was asked about the commonalities between blood and water. His response was: “What’s alike about them is that water washes off all sorts of dirt, so it can wash off blood too!” (p. 82). The response of another peasant who was asked about the commonalities between a fish and a crow was: “If the fish just lays on top of the water, the crow could pinch it!” (p. 82). However, our subjects were neither young nor illiterate. Hence, their performance cannot be readily explained by insufficient cognitive or cultural development. Moreover, their performance cannot be explained away by misinterpretation of the task at hand. In fact, 214 of the 234 scenarios generated by 73% of the very bright undergraduates who participated in our study were written next to high-quality analogical arguments that explained the similarity between the base and the relational targets in the same triplets. How then can one explain the systematicity and prevalence of scenario justifications generated by our subjects?

Let us first consider two explanations that cannot account for our results. One possibility is to argue that the scenario explanations generated by our subjects were the result of uncontrolled interference from highly activated thematic associations that erroneously replaced the task-appropriate but less salient comparison inference. However, the nature of our stimuli and the creative scenarios generated by our subjects do not lend support to the interference explanation. First, the attributional targets were not designed to make the comparison process especially difficult. On the contrary, these targets had two salient matches with the base (i.e., the two nouns) that could be used to explain similarity and one salient mismatch (i.e., the verb) that could be used to explain difference. Second, the attributional targets were not designed to induce salient thematic associations. Even in hindsight, these targets do not appear to be obviously slanted toward familiar associations with the base statements. For example, there is no obvious causal or temporal relation between catching a ball and selling it or between accuracy of equations and their difficulty. In fact, our subjects often used this lack of thematic relatedness to explain why the paired statements were not similar (i.e., impeding scenarios). Thus, it is unlikely that the attributional targets caused thematic interference. Rather, it appears that the attributional targets induced an active search for possible thematic relations that turned out to be either successful (i.e., supportive scenarios) or unsuccessful (i.e., impeding scenarios).

The nature of our stimuli also makes it unlikely that the impact of thematic scenarios can be incorporated into traditional models of similarity by treating a joint scenario, or lack of it, as an inferred abstract attribute that either increases or decreases similarity. Again, this play may work for well-established scripts and scenarios (cakes and presents being part of a birthday party script), but it ignores the creative and constructive aspects of thematic integration in the scenarios generated by our subjects. If people were asked to list features of birthday parties, they no doubt would mention both presents and a cake, but it is quite unlikely that they would spontaneously mention that carpenters sit in chairs as part of a chair-fixing scenario or that pianists organize competitions in order to secure a win. Importantly, this argument cannot account for thematic scenarios that were generated as countereftuals to explain lack of similarity (e.g., “It does not necessarily follow that because he fixed the chair he had to sit in it”!). Instead, it appears that comparison and integration are two distinct processing strategies that mediate similarity judgments. At present, we do not know whether comparison and integration operate simultaneously to support and impede similarity, whether they compete with each other, or whether integration comes into play only when comparison fails to achieve some satisfying criteria. Nevertheless, we are quite con-
fident that the relative prominence of compari-
son and integration depends on the alignability of
the paired stimuli. The process of compari-
sion demands that the representations of the
paired stimuli will be aligned. Indeed, several
studies document that people are quite cre-
ative in their attempts to achieve alignment.
For example, they may interpret an ambiguous
attribute of a target stimulus such that it will
fit an attribute of a given base (e.g., Medin et
al., 1993), or may even try to invent novel
alignable attributes (e.g., Markman & Medin,
1995). However, certain discrepancies be-
tween the paired stimuli may be difficult to
overcome by such creative acts of interpreta-
tion. When this happens, the process of the-
matic integration becomes prominent. That is,
we believe that difficulties in aligning the
paired stimuli set the ground for the emer-
gence or sufficient activation of the integration
process.

This account builds upon and sets the lim-
it ing conditions of the structural-alignment ap-
proach to similarity. Specifically, it appears
that purely attributional matches (e.g., match-
ing nouns) are insufficient for a satisfactory
alignment of structured stimuli (e.g., combina-
tions of interrelated objects). Hence, even
though the nouns in the attributional targets
were obviously alignable (e.g., “Both involve
a carpenter and a chair”), structural alignment
was severely impaired by mismatching verbs
(e.g., “Sit differs from fix”) allowing the-
matic integration to win over and become the
more prominent strategy. The alignability ac-
count also implies that when the paired stimuli
are objects (rather than structured combina-
tions of objects), thematic integration should
be more prominent when the paired objects
are nonalignable (i.e., few common attributes)
when they are alignable (i.e., many com-
mon attributes). Witemski and Bassok
(1996) tested this hypothesis in a recent fol-
low-up study. They found that, indeed, the-
matic scenarios were much more frequent in
similarity judgments of nonalignable object
pairs (e.g., milk-horse) than in judgments of
alignable object pairs (e.g., milk-lemonade).

Salient preexisting scenarios simply magni-
ﬁed this differential tendency. Speciﬁcally,
subjects generated more thematic scenarios
and gave higher similarity ratings to themati-
cally related than to thematically unrelated
pairs of objects that were equated in alignabil-
ity (e.g., milk-cow > milk-horse; milk-coffee
> milk-lemonade).

Future research will have to establish
whether and how the alignability of the paired
stimuli affects the interplay between compari-
sion and thematic integration in similarity
judgments. Needless to say, the relative prom-
inence of these two processing strategies is
likely to be affected by enculturation (e.g.,
Luria, 1976), maturity (e.g., Markman, 1989),
and task instructions. For example, it is quite
reasonable to assume that thematic scenarios
will be more likely to occur when people are
not repeatedly reminded of the colloquial com-
parison meaning of similarity (e.g., hav-
ing to judge relational targets in the same trip-
let) when they are not made accountable for
their reasoning process (i.e., explanations) and
when they have little time to arrive at their
judgments. Thus, it appears that our study pro-
vided very unfavorable conditions for them-
atic replacements. Yet, at this point, we do
not know whether, how, and to what extent
these various pressures affected the tendency
of our subjects to rely on thematic integration
while judging similarity. On the one hand, the
above analysis suggests that mature adults in
Western culture may try to actively “undo”
the impact of thematic scenarios on their simi-
larity ratings. If so, then our results may un-
derestimate the frequency of thematic integra-
tion in normal processing. On the other hand,
one could focus on the observation that the
magnitude of attributional ratings in Experi-
ment 3 (explanations) was somewhat higher
than in Experiment 2 (no explanations) and
argue that explanations actually increased sub-
jects’ natural tendency to integrate the paired
stimuli into coherent scenarios. Future re-
search will have to establish whether thematic
scenarios are as prevalent as is suggested by
our results.
Of course, one could define these results out of existence by restricting "similarity" to denote only comparison and argue or acknowledge that another process (e.g., "association") frequently gets in the way of similarity judgments. Although this rather high-handed approach fosters panimory, it ignores the fact that a single processing mechanism cannot account for the way in which people construe similarity.

The Ubiquity of Processing Replacements

Theoretical accounts of many cognitive processes posit that similarity provides the input to, or is used in lieu of, other inferential mechanisms. Such similarity-based accounts also posit that similarity reflects the degree of correspondence between the properties of the paired entities. However, there is substantial evidence that people's performance in such tasks may be also mediated by thematic integration. For example, the representativeness heuristic by which people tend to estimate the probability of uncertain events (Kahneman & Tversky, 1972) has been often equated with similarity (Smith & Osherson, 1989). However, according to Tversky and Kahneman (1983), representativeness is a broader term that can denote both the degree of correspondence between an instance and a category (e.g., a shy person wearing glasses and a prototypical librarian) and the relation between an act and an actor (e.g., the act of dancing and a slim woman dressed in a leotard). That is, probability judgments are sometimes mediated by similarity and sometimes by thematic associations.

Another example is the process of categorization which, in most cases, is believed to be mediated by similarity between an instance and the category representations (e.g., Estes, 1994). Yet, people also form ad hoc categories (Barralou, 1963, 1991) on the basis of thematic relations (e.g., grouping a toothbrush with a wall-lamp because both are things one should take when going on a trip). More recently, Lin (1996) has shown that college students, much like young children, have a strong preference to form categories that include thematically related entities (e.g., cat and litter box) over categories that include entities from the same taxonomic category (e.g., cat and lion) and to use thematically related categories in reasoning. Finally, mass superordinate categories (e.g., furniture) appear to be structured not so much in terms of alignable properties as in terms of thematic relations (Murphy & Wason, 1989).

The revised direction of processing replacements has been documented by Wisniewski (1996) who identified two strategies by which people combine pairs of concepts that correspond to the two processing strategies we identified in similarity judgments. Specifically, he found that although most concepts are combined by a thematic link (e.g., shoe-box is understood as a box that contains shoes), when the two concepts are similar (e.g., skunk-squirrel) they are combined by integrating their respective attributes (e.g., a bad smelling squirrel). Bussok and Chase (1996) found that a similar distinction between alignable and nonalignable objects guides people's reasoning about arithmetic word problems. For example, when asked to construct simple division problems involving various pairs of object sets, people readily divided thematically related objects (e.g., apples-baskets) but refrained from dividing similar objects (e.g., apples-oranges). Instead, they constructed more complex problems in which the similar objects were added to each other (e.g., apples + oranges)(children).

The above examples demonstrate that the use of thematic associations revealed in the justifications of our subjects is not a unique and exceptional phenomenon. Rather, partition and integration seem to mediate performance on a variety of tasks, emerging (spontaneously) as the prominent processing strategies for alignable and nonalignable stimuli. We believe that such sensitivity to the nature of the stimuli is likely to result in more meaningful inferences. For example, forming a causal relation between the accuracy and the difficulty of equations may lead to more
meaningful inferences than identifying commonalities and differences between such non-alignable concepts (e.g., Markman & Gentner, 1993). Importantly, these examples demonstrate that the boundaries between various inferential processes are not as sharp as might be implied by the distinct labels given to them in natural language and in theoretical accounts of cognition.

The Constructive Nature of Similarity

The processing replacements documented in our study are an extreme example of the constructive nature of similarity (Medin et al., 1999), much more extreme than any changes in weights assigned to relational and attributional matches in response to changes in task instructions (e.g., similarity vs difference; explanations vs no explanations). At the same time, our results provide an unexpected support for the psychological validity of the alignment hypothesis and for the distinction between relations and attributes in similarity judgments. Not only do people rely on this distinction when they align and compare the representations of the paired stimuli (e.g., ratings of the relational targets), they actually use this distinction to decide whether to compare or to integrate the paired stimuli (i.e., relational vs attributional targets, respectively).

In the present paper we were mainly concerned with the possibility of identifying processing regularities. Needless to say, the constructive nature of similarity judgments was extremely salient at the stage of developing representations—when people were using their semantic knowledge to interpret the noun-verb noun statements. Matches in the combined meanings of the base and target statements included well-reasoned arguments about the truth, plausibility, and informativeness of these statements (e.g., "The fact that cottage cheese is cheaper than a house doesn’t say anything we don’t already know"). Such well-reasoned elaborations of the statements also appeared in the thematic scenarios (e.g., "A cathedral might be higher than a church because it is older, since people used to build higher buildings in older times"). That is, the constructive nature of similarity was apparent in people’s extensive reliance on intuitive and deductive inferences by which they exploited their semantic knowledge to interpret the base and target stimuli. These sense-making inferences and heavy reliance on knowledge of nonarbitrary dependencies across various aspects of the stimuli have major effects on performance (see Murphy & Allopenna, 1995, and Spaulding & Murphy, 1996, for similar findings obtained in the context of concept learning). Yet, such inferences remain outside the scope of processing models that attempt to explain similarity or performance that is mediated by similarity. Importantly, it is typically assumed that because similarity (and association) is a more basic and a more automatic processes than is rule-based reasoning, it often supports, impedes, or replaces the more complex inferences (see Stolman, 1996 for a review). The constructive processes by which our subjects exploited their semantic knowledge demonstrate that the relation between similarity and these more complex inductive and deductive inference rules might be bidirectional. That is, rule-based reasoning might support, impede, or even replace similarity judgments (and associations). A related example is the study of Chase and Easson (1995) who found that people apply formal probability rules to compute similarity between an instance and a conjoint category, reversing the direction of influence between similarity and probability judgments (i.e., the representativeness heuristic).

Of course, when the stimuli consist of arbitrary combinations of attributes and relations (e.g., interrelated geometric figures), the impact of such inductive inferences is substantially reduced. For instance, it is unlikely that people will attempt to interpret the truth or informativeness of the analogy between "A triangle is above a square" and "A star is above a circle" or will attempt to construct a causal scenario to determine the similarity between "A triangle is above a square" and
"A triangle is inside a square." Although studies that use such stimuli are less "contaminated" by semantic inferences, they underestimate the constructive nature of similarity judgments and might fail to discover ecologically valid processing regularities (e.g., thematic integration). We believe that the advantage of finding such regularities justifies the difficulty of sorting through the richness and complexity of the intervening inferential processes.

Appendix
Quintuplets of Base and Target Statements Used in Experiments 1 through 3

I. Action Statements
B1: The carpenter fixed the chair.
AA: The carpenter sat on the chair.
RR*: The electrician fixed the radio.
R: The plumber fixed the radio.
RA: The carpenter fixed the radio.
B2: The teacher prepared the lecture.
AA: The teacher listened to the lecture.
RR*: The cook prepared the meal.
R: The tailor prepared the meal.
RA: The teacher prepared the meal.
B3: The engineer designed a car.
AA: The engineer drove a car.
RR*: The choreographer designed a dance.
R: The lawyer designed a dance.
RA: The engineer designed a dance.
B4: The pianist won the competition.
AA: The pianist organized the competition.
RR*: The horse won the race.
R: The violinist won the race.
RA: The pianist won the race.
B5: The barber cut the hair.
AA: The barber thought about the hair.
RR*: The teenager cut the lawn.
R: The insurance agent cut the lawn.
RA: The barber cut the lawn.
B6: The lawyer took the case.
AA: The lawyer examined the case.
RR*: The teacher took the position.
R: The thief took the position.
RA: The lawyer took the position.
B7: The child enjoyed the toy.
AA: The child hid the toy.
RR*: The woman enjoyed the job.
R: The librarian enjoyed the job.
RA: The child enjoyed the job.
B8: The boy caught the ball.
AA: The boy sold the ball.
RR*: The angler caught the fish.
R: The news reporter caught the fish.
RA: The boy caught the fish.

II. Comparison Statements
B9: The cathedral is higher than the church.
AA: The cathedral is older than the church.
RR*: The tree is higher than the bush.
R: The house is higher than the bush.
RA: The cathedral is higher than the bush.
B10: The tenor sang nicer than the soprano.
AA: The tenor sang lower than the soprano.
RR*: The tricklingbird sang nicer than the crow.
R: The dancer sang nicer than the crow.
RA: The tenor sang nicer than the crow.
B11: Cottage cheese is cheaper than mozzarella cheese.
AA: Cottage cheese is healthier than mozzarella cheese.
RR*: A tent is cheaper than a house.
R: A book is cheaper than a house.
RA: Cottage cheese is cheaper than a house.

B12: Equations are more accurate than words.

AA: Equations are more difficult than words.

RR*: Missiles are more accurate than bombs.

R: Computers are more accurate than bombs.

A: Equations are more accurate than bombs.

REFERENCES


ments of similarity and difference are not inverse. Psychological Science, 1, 64–69.


(Received February 21, 1996)

(Revision received July 9, 1996)