

# Structural alignment facilitates the noticing of differences

DEDRE GENTNER and VIRGINIA GUNN  
*Northwestern University, Evanston, Illinois*

High-similarity concept pairs that elicit many commonalities also elicit many related differences (Gentner & Markman, 1994; A. B. Markman & Gentner, 1993a, 1993b, 1996; A. B. Markman & Wisniewski, 1997). This finding has been used to support the claim that the comparison process is one of structural alignment. However, it is possible that the difference advantage results from some other property of high-similarity pairs, such as a greater number of stored differences. The present experiments demonstrate that the comparison process itself leads to the greater psychological availability of differences. In three experiments, participants listed commonalities for word pairs and then listed differences under a time pressure for these old pairs and new pairs. In Experiment 1, participants listed more differences for old than for new pairs, consistent with the claim that the comparison process facilitates noticing differences. In Experiment 2, we showed that the difference-listing advantage is specific to the comparison process: Mere coprocessing of the pairs (specifically, providing thematic relations) does not facilitate, and in fact appears to inhibit, difference listing. In Experiment 3, pairs with deeper common systems elicited a larger number of specific alignable differences than did pairs with shallow sets of commonalities. Overall, the results support the structural alignment claim that the comparison process promotes the noticing of both commonalities and related differences.

The perception of likeness is practically very much bound up with that of difference. That is to say, the only differences we note as differences, and estimate quantitatively, and arrange along a scale, are those comparatively limited differences which we find between members of a common genus.  
(James, 1890, Vol. 1, p. 528)

How are differences generated and why are certain differences noticed and others are not? The answer may depend on whether or not differences bear any relationship to commonalities. According to *mental distance* models of similarity (Nosofsky, 1987; Shepard, 1974; Shoben, 1983), the degree of difference is the inverse of the degree of similarity on any given dimension. In *independent feature* models (e.g., Tversky, 1977), differences are independent of commonalities; the differences between a pair of objects are simply any elements of the objects' feature sets that do not match. Although commonalities and differences may be differentially weighted according to the task or context, there is no necessary relationship between the common features and the psychologically salient differences.

In contrast to the above accounts, the *structural alignment* approach (Gentner, 1983; A. B. Markman & Gentner, 1993a, 1993b; Medin, Goldstone, & Gentner, 1993) posits representations composed of interconnected structures, rather than independent features or dimensional spaces. According to this view, as discussed below, differences are noticed relative to commonalities—that is, one first notices commonalities and then the differences that are related to those commonalities (e.g., whales and fish are both swimming creatures, but one has lungs, the other has gills).

In the present series of experiments, we demonstrate four findings that link the structural alignment process with the generation of differences. First, differences are easier to generate for word pairs that have been recently aligned than for those that have not, demonstrating that structural alignment facilitates noticing differences. Second, this facilitation is specifically related to structural alignment and not merely the result of joint activation of the word pairs. Third, the degree of difference facilitation reflects the quality and extent of the pairs' common system. Fourth, pairs with deep and rich alignments elicit not only more differences, but more differences specifically related to the commonalities than do pairs with more sparse alignments.

Structural alignment theory is a generalization of the structure-mapping theory of analogical reasoning (Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1983), according to which comparison is accomplished by a process of alignment of structured representations of the entities or scenes being compared and the subsequent projection of inferences (Gentner & Markman, 1994, 1997; Goldstone, Medin, & Gentner, 1991; A. B. Markman &

---

This work was supported by an NSF Graduate Fellowship awarded to V.G. and by NSF Grant SBR-95-11757 and ONR Grant N00014-92-J-1098 awarded to D.G. This paper was partially prepared while D.G. was a fellow at the Center for Advanced Study in the Behavioral Sciences. We are grateful for the financial support provided by the William T. Grant Foundation, Award 95167795. We thank Doug Medin, Ken Kurtz, Art Markman, and the entire Similarity and Analogy Group at Northwestern University for helpful discussions. Correspondence should be addressed to D. Gentner, Psychology Department, Swift Hall, Northwestern University, Evanston, IL 60608 (e-mail: gentner@northwestern.edu).

Gentner, 1993a, 1996, 1997; Medin, Goldstone, & Gentner, 1990). The process of structure-mapping attempts to place two representations in correspondence so that they form the maximal (i.e., the largest and deepest) globally consistent match. This match must be structurally consistent—that is, it must satisfy two constraints: *parallel connectivity*, which requires that arguments of matching predicates must themselves be able to be placed in correspondence, and *one-to-one correspondence*, which requires that each element of a representation match, at most, one element of the other representation. Finally, people prefer deep matching systems over those with only isolated, scattered matches (the *systematicity principle*) and draw inferences based on completing systematic patterns. (Clement & Gentner, 1991; Gentner & Rattermann, 1991; A. B. Markman & Gentner, 1993b).

Once the maximal system is found, a further hypothesis is that differences associated with the system (e.g., different objects in the same role, or different attributes of corresponding objects) become salient. Thus, pairs with a larger system of commonalities have more potential for generating salient differences. This leads to the rather surprising prediction that differences should be easier to generate for high-similarity pairs (e.g., *bicycle–tricycle*) than for low-similarity pairs (e.g., *navy–sculpture*). This prediction is unique to the structural alignment account. The *mental distance* account makes no clear predictions, and the *independent feature* account predicts that, all else being equal, differences should be easier to generate for low-similarity pairs than for high-similarity pairs.

Previous research has provided considerable evidence that object pairs with many commonalities also elicit many related differences (Gentner & Markman, 1994; A. B. Markman & Gentner, 1993a, 1993b, 1996). This work has largely focused on the distinction between “alignable” and “nonalignable” differences. An alignable (related) difference is one that is conceptually related to a commonality, whereas a nonalignable difference is not conceptually related to a commonality. For example, the fact that a bicycle has two wheels and a tricycle has three wheels is an alignable difference related to the commonality that both are wheeled vehicles. In contrast, the fact that a bicycle has handlebars, whereas a refrigerator does not is a nonalignable difference. This distinction has often been operationalized in terms of a correlated response tendency: Alignable differences are typically stated as explicitly differing values on a common dimension or predicate (e.g., *squirrels have fluffy tails, mice have thin tails*), whereas nonalignable differences are stated by asserting a fact for one item and denying it for the other (*squirrels have feet, carpets don't*).

There is evidence that alignable differences are related to commonalities. For example, A. B. Markman and Gentner (1993a) found that when participants listed commonalities and differences for high-similarity and low-similarity pairs of words the number of alignable differences was positively correlated with the number of

commonalities, whereas the number of nonalignable differences was negatively correlated with commonalities. Further, when new participants were given these previously generated commonalities and differences and asked to match any that were conceptually related, they sorted alignable differences, but not nonalignable differences, with the commonalities. Similar results were obtained with pictorial scenes (A. B. Markman & Gentner, 1996).

These studies show that alignable differences increase with commonalities.<sup>1</sup> Support for the stronger claim that psychologically noticeable differences, in general, increase with commonalities comes from a *speeded difference* task, which provides a measure of the psychological availability of differences (Gentner & Markman, 1994). Participants were given a large set of word pairs—half high-similarity and half low-similarity—and asked to list a single difference for as many pairs as possible in a limited time period. According to structural alignment theory, high-similarity pairs should show an advantage because (1) they should be easy to align and (2) the alignment process should yield a large common system from which to derive related differences. The results were striking: Differences were listed for over twice as many high-similarity pairs as for low-similarity pairs.

Gentner and Markman (1994) interpreted this high-similarity advantage as support for structure-mapping theory. However, it is possible that this difference-listing advantage for high-similarity pairs could simply reflect a disparity in the number of prestored differences due to past experience. People are more likely to have had occasion to compare high-similarity pairs than low-similarity pairs—for instance, to have weighed the relative advantages of a hotel and a motel. Any such precomputed differences would presumably come to mind easily. One argument against this concern is A. B. Markman and Gentner's (1996) finding of similar patterns of difference responses for high- and low-similarity picture pairs, for which prestored differences seem unlikely. But a definitive test of this competing explanation requires a more direct test of the effects of structural alignment.

If indeed it is the structural alignment *process* that generates related differences, it should be easier to list differences for word pairs that have been recently compared than for word pairs that have not. Thus, in this research, we had the participants list commonalities for a set of pairs—thus inducing alignment processing—and then we assessed whether their ability to give differences for these pairs was elevated with respect to neutral pairs. For this purpose, we wanted to use a theory-neutral measure of difference facilitation rather than invoke the alignable/nonalignable distinction. Therefore, we used the Gentner and Markman (1994) speeded difference task and simply counted the overall number of differences produced.

The participants completed a commonality listing task for several pairs of both high and low similarity, followed by a speeded difference task on the same pairs and on an equal number of new pairs. As was discussed above, the

speeded difference task requires participants to write one difference for as many pairs as possible in a limited time period. The first prediction is that the participants will list a difference for more high-similarity pairs than low-similarity pairs (similarity main effect). If only this effect is obtained, the advantage for high-similarity pairs may be accounted for by prior stored differences. However, the main prediction of interest is that the participants will list a difference for more of the old pairs than new pairs (experience main effect). Such a result would show that prestored differences are not the whole story and that the recently completed alignment process also contributes. (The possibility that any such benefits might derive simply from activating the pairs is addressed in Experiment 2.) Lastly, although the participants should generate many more differences for high- than for low-similarity pairs, the increase in differences for old pairs over new pairs may be more pronounced for the low-similarity pairs than for the high-similarity pairs (manifested as an interaction). Since the high-similarity pairs have large and consistent common systems, they should be easy to align, even when seen for the first time. This assumption was verified in a pilot speeded commonality task. When asked to list a single commonality for as many of 40 pairs as they could within 4 min, the participants ( $n = 8$ ) listed a commonality for more of the high-similarity pairs ( $M = 17.9$ ) than the low-similarity pairs ( $M = 5.3$ ) [ $F(1,7) = 167.23$ ,  $MS_e = 4.0$ ,  $p < .0001$ ]. If indeed low-similarity pairs are less readily compared spontaneously than high-similarity pairs, then their comparison likelihood may be more strongly influenced by instructions to compare.

**EXPERIMENT 1**

**Method**

**Participants.** Forty-eight paid participants were recruited from the Northwestern University community.

**Materials and Design.** The materials, listed in the Appendix, were based on the word pairs used by Gentner and Markman (1994). As in Gentner and Markman’s original study, the low-similarity pairs were constructed by re-pairing the words used in the high-similarity list, thus ensuring that any differences between high and low similarity could not be attributed to the words themselves. (A few low-similarity pairs were re-paired for this study to in order to eliminate incidental commonalities that appeared in pilot testing.) The materials for the speeded difference test were two stimulus sets of 40 pairs, each presented in one random order and its reverse. In the initial commonality task, each group of participants received one of these four subsets, presented either in one random order or in its reverse. In the subsequent speeded difference task, they again received these 20 pairs plus another subset of 20 new pairs. The design was stimulus subset (four levels, between-subjects)  $\times$  similarity type (high or low, within-subjects)  $\times$  experience (old or new pairs, within-subjects).

**Procedure.** The participants first listed a single commonality for each of 20 word pairs (half high similarity and half low similarity). The participants were not told about the task to follow. There was no time limit, but the participants generally completed this task in 10 min.

The participants then performed a speeded difference task. They were told to list a single difference for as many of the pairs as possible in 5 min. They were encouraged to complete the “easiest” pairs first, since they would not have enough time to respond to all of the

pairs. For this task, the participants received 40 pairs—20 old (i.e., compared in the commonality task) and 20 new. Half of the pairs were high similarity and half were low similarity. All word pairs were listed on a single page, and the participants were timed with a stopwatch.

**Results and Discussion**

Sample responses are given in Table 1. As expected, the participants produced a difference for more of the high-similarity pairs ( $M = 6.86$ ) than low-similarity pairs ( $M = 4.39$ ) [ $F_s(1,47) = 36.69$ ,  $MS_e = 8.04$ ,  $p < .001$ ]. The key prediction of an effect of experience was borne out: The participants produced differences for more old (previously aligned) pairs ( $M = 5.87$ ) than new pairs ( $M = 5.37$ ) [ $F_s(1,47) = 4.51$ ,  $MS_e = 2.66$ ,  $p < .05$ ]. However, contrary to expectation, the effect of prior alignment was not greater for low-similarity pairs ( $M_{new} = 4.23$ ;  $M_{old} = 4.52$ ) than for high-similarity pairs ( $M_{new} = 6.52$ ;  $M_{old} = 7.19$ ). There was no interaction between similarity and experience nor any main effect or interactions with the grouping variable (stimulus subsets) in the subject analysis. An item analysis also revealed an effect of similarity type [ $F_i(1,72) = 104.53$ ,  $MS_e = 3.30$ ,  $p < .001$ ] and experience [ $F_i(1,72) = 7.96$ ,  $MS_e = 1.59$ ,  $p < .01$ ]. There was a three-way interaction: The interaction between similarity and experience differed across stimulus sets [ $F_i(3,72) = 6.69$ ,  $MS_e = 1.59$ ,  $p < .01$ ].

The results show that the alignment process facilitates noticing differences. The word pairs that had been previously aligned were the same pairs that tended to elicit differences. Further, many of these word pairs—particularly the low-similarity pairs—were unlikely to have had prestored differences available. The superiority of high-similarity pairs could indicate the presence of prestored differences, but could also have arisen from the greater ease of aligning high-similarity items *de novo*. However, the prestored differences explanation cannot account for why the *same* pairs elicited differences more often when they were old than when they were new (the *experience effect*). This finding suggests that differences can be generated on line following alignment experience.

**Table 1**  
**Sample Responses From Experiment 1**

<i>Football–Hockey</i> (High Similarity)	<i>Sculpture–Navy</i> (Low Similarity)
<b>Distinct commonalities listed</b>	<b>Distinct commonalities listed</b>
Sports	Mysterious, hard to understand
Both involve physical contact	Both found all over world
Wear padding for both	Both are decorated
Both have team bench areas	Both can be still
	Have lots of metal
	Often near water
	Both interesting to look at
<b>Distinct differences listed</b>	<b>Distinct differences listed</b>
Shape of ball/puck; skates vs. cleats	One is artistic, made by artist
Played in different seasons	Art vs. defense
Ball vs. puck	Refined vs. crude
Football played on 100-yard field	Navy has people
Different rules	
One on grass/one on ice	
Warmer/cooler	
One uses feet/the other, sticks	

## EXPERIMENT 2

The results of Experiment 1 provide evidence that carrying out a structural alignment process facilitates noticing differences. These results are potentially strong support for structural alignment theory and argue against theories of similarity in which commonalities and differences are independent. However, a weaker explanation must be considered. In the commonality task, both concept representations are accessed together in one processing context. Perhaps this coactivation potentially accounts for the old-pair advantage in the difference task.

The question is whether difference facilitation is particular to structural alignment or whether it would occur following any type of concurrent activation of word pairs. In order to find out, we examined the performance of two groups of participants in a speeded difference task. One of the groups performed a prior commonality listing task, as in Experiment 1. The other group completed a thematic connection task with the same word pairs before listing their differences. A thematic connection between *babysitter* and *phone*, for example, might be *a babysitter often talks on the phone*. The thematic relations task was chosen because it requires a meaningful coactivation of the pairs but does not entail a structural alignment. The contrast between the two groups in the difference-listing task should reveal whether the structural alignment process contributes anything over and above mere coactivation of concept representations.

There was a second reason for choosing the thematic relations task. Some recent results have suggested that thematic association and similarity may be subcomponents of a single mental process. Spontaneous listings of thematic associations have been observed in tasks that simply asked for commonality and difference listings (A. B. Markman & Gentner, 1993a) or justifications of similarity ratings (Bassok & Medin, 1997). Further, thematic relations between items can increase their judged similarity, so that *milk-coffee* is rated more similar than *milk-lemonade* (Wisniewski & Bassok, 1996). Such findings might suggest that similarity and thematic associations are closely related (see Sloman, 1996). However, it could be that thematic associations merely intrude on the comparison process and hence should not be thought of as a legitimate aspect of similarity (Gentner & Brem, 1998). If carrying out a thematic task induces different performance on the difference task than carrying out a comparison task, this will suggest that the two processes are distinct (though both may feed into a general feeling of relatedness). Indeed, in this case the thematic group might even be expected to generate fewer differences for the old pairs than for the new pairs, if focusing on thematic relations competes with the alignment process.

The predictions are as follows. Structural alignment theory predicts that the commonality group, but not the thematic group, should list a difference for more of the old pairs than the new pairs, because these participants will have structurally aligned the old pairs in the first task.

If coactivation is sufficient, or if similarity and thematic association are the same process, then *both* groups will demonstrate difference facilitation.

### Method

**Participants.** Ninety-six Northwestern University undergraduates participated. The results of one participant, who failed to follow directions, were excluded, leaving 95 participants.

**Materials and Design.** The materials are listed in Table 2. We used low-similarity word pairs that could yield either commonalities or thematic connections. (Low-similarity pairs were used because high-similarity pairs might have invited alignment even in the thematic-relations task, thus undermining the design.) For example, the pair *tree-child* could elicit the commonality *both grow* or the thematic connection *a child climbs a tree*. Likewise, the word pair *cult-FBI* could elicit either *both are secretive* or *the FBI investigates cults*. The same word pairs were used in the thematic connection and commonality tasks.<sup>2</sup>

In the initial setting task—commonality or thematic connection—the participants saw one of two stimulus sets of 20 word pairs. The speeded difference task that followed consisted of the full set of 40 word pairs. The word pairs that were “old” for one participant group were “new” for the other. The materials were presented in 24 different random orders. The design was stimulus subset (two levels, between-subjects)  $\times$  instruction task (thematic connection or commonality task, between-subjects)  $\times$  experience (old or new pairs, within-subjects).

In both the thematic connection and the commonality setting tasks, five additional anchoring pairs were distributed evenly among the others to encourage the desired mode of processing. For example, the pairs *knife-butter* and *sock-foot*, provided only on the thematic connection task, were chosen as having strong salient thematic connections. Likewise, the pairs *tadpole-minnow* and *blinds-sunglasses* were included on the commonality task because they have strong commonalities. None of the anchoring pairs appeared on the actual difference task.

**Procedure.** The participants were randomly assigned to either the thematic group, which listed a single thematic connection for each of the word pairs, or the commonality group, which listed a single commonality for each word pair. As in Experiment 1, there was no time limit for this initial task, and the participants were not told that a related task would follow. After the initial task, the participants completed a speeded difference task with 40 pairs, half old and half new.

### Results and Discussion

Sample responses are given in Table 3. The commonality group produced more differences overall ( $M = 8.77$ ) than did the thematic group ( $M = 7.24$ ) [ $F_s(1,91) = 5.92$ ,  $MS_e = 18.68$ ,  $p < .05$ ]. More importantly, as predicted, the setting task experience was beneficial only for the commonality group [ $F_s(1,91) = 9.27$ ,  $MS_e = 5.57$ ,  $p < .01$ ]. The commonality group produced differences for more old pairs ( $M = 9.34$ ) than new pairs ( $M = 8.2$ ) [ $t(46) = 2.04$ ,  $p < .05$ ]. As predicted, the thematic group did not show this difference facilitation ( $M_{old} = 6.77$ ;  $M_{new} = 7.71$ ), as is shown in Figure 1. In fact, the thematic group performed worse with old pairs than with new pairs [ $t(46) = 2.04$ ,  $p < .05$ ]. Experience with the old pairs appeared to hamper the process of finding differences. The number of differences listed in each instruction condition also varied with stimulus set [ $F_s(1,91) = 5.11$ ,  $MS_e = 18.68$ ,  $p < .05$ ]. An item analysis confirmed the main ef-

**Table 2**  
**Word Pairs Used in Experiment 2**

Camera	Lighthouse	Tree	Child
Mask	Wig	Training wheels	Garage
Blanket	Tent	Soft drinks	Television
Fence	Floodlights	Anthill	Poison
Refrigerator	Tupperware	Doily	Antique table
Catalog	Calendar	Pager	Chain
Ice	Skylight	Surgeon	Pianist
Locket	Safe deposit box	Grate	Tunnel
Mind	Factory	Car alarm	Radio
Library	Attic	Shoe	Tire
Deer hunter	Mosquito	Photo	Class ring
Key	Bribe	Action movie	Earthquake
Life boat	Fire extinguisher	Shovel	Roots
Relic	Gold	Cult	FBI
Fireplace	Comforter	Gas mask	Quarantine
Stars	Map	Magnifying glass	Microphone
Killer whale	Shark	Tightrope	Unicycle
Lungs	Bagpipe	Blueprint	Robbery
Yacht	Mop	Travel game	Briefcase
Ladder	Chimney	Pothole	Storm
Anchors for Commonality Task		Anchors for Thematic Task	
Tadpole	Minnow	Knife	Butter
Socket	Electric eel	Sock	Foot
Wool	Blubber	Cornflakes	Spoon
Blinds	Sunglasses	Skateboard	Sidewalk
Battery	Coal	Dog	Stick

Note—All word pairs, except for the anchoring pairs, were designed to participate in either a commonality or a thematic relation and were given to both groups of participants. Anchoring stimuli did not appear on the difference task for either group.

fect of instruction condition [ $F_1(1,38) = 30.4, MS_e = 3.26, p < .01$ ] and the interaction between instruction condition and experience level [ $F_1(1,38) = 4.58, MS_e = 3.15, p < .05$ ]. There was also a three-way interaction between instruction condition, experience level, and stimulus set [ $F_1(1,38) = 18.32, MS_e = 3.15, p < .01$ ].

Interestingly, the participants occasionally confused commonalities and thematic relations in Experiment 2, listing a thematic relation when asked for a commonality or vice versa. However, the degree of difference facilitation was related to the total number of commonalities that were produced, regardless of instructions ( $r = .31, p < .05$ ). In contrast, the observed difference facilitation was not related to the number of thematic relations produced, regardless of instructions ( $r = -.19, p > .2$ ). This supports the claim that salient differences arise from the common alignment.

**Table 3**  
**Sample Responses from Experiment 2 for the Word Pair: Locket–Safe Deposit Box**

Commonality Subjects	Thematic Subjects
<b>Consensus commonalities</b> Both close, lock for privacy Keep things of value safe	<b>Consensus thematic connection</b> Locket may be placed in a safe deposit box
<b>Consensus difference</b> One holds things of emotional Value; the other holds things Of financial value	<b>Consensus difference</b> [no consensus] One larger One more expensive You can wear one

The above evidence supports the assertion that it is the process of structural alignment specifically, and not simply joint consideration of the terms, that facilitates the generation of differences. The participants who listed a thematic connection between words produced fewer differences for the pairs they had seen before than for the pairs that were new to them. It seems that mere coprocessing of the word pairs is not sufficient to produce difference facilitation, and indeed, some types of processing may even interfere with the generation of differences. These results suggest that facilitation of differences results specifically from structural alignment processes.

**Relational and Attributional Similarity: The Quality of Alignment Matters**

The evidence implicates the comparison process in difference generation. Can we push this connection further? According to the structure-mapping account, the larger and deeper the interconnected structure, the easier it should be to generate related differences. One way to test this prediction is to contrast relationally similar and attributionally similar pairs. As mentioned earlier, relations take two or more arguments (entities or other relations), whereas attributes take a single argument. Thus, relational matches—even first-order relational matches with only object arguments—provide a larger common system from which to derive differences than do attributional matches. Further, relational matches are deeper than attribute matches in the following sense: Attributes function as object descriptors (and are thus dominated by

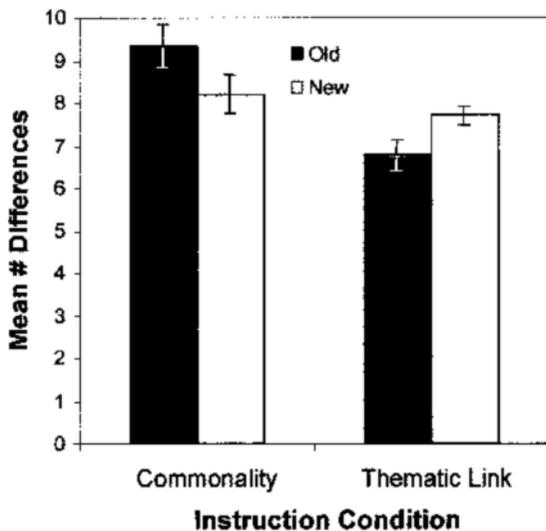


Figure 1. Results of Experiment 2: Mean differences  $\times$  instruction condition and experience.

their objects), whereas relations create a relational structure *between* objects. Thus, extracting relational correspondences should be more useful in subsequently finding differences than should extracting attributional matches.

For example, the pair *tiger–shark* could yield the relational commonality *both pursue other creatures in order to eat them*. From this may stem many connected differences (e.g., type of prey, hunting behaviors, etc.). In contrast, a salient attributional commonality—for example, that a pancake and a nickel are circular—invites few related differences, because this commonality is not well connected to other concepts within the common structure. Consider also the low-similarity pairs *bowl–dictionary* and *watch–banana split* from Experiment 1. The most frequently listed commonality for *bowl–dictionary* was *both contain things*. Five of the 7 participants who found this (relational) commonality also found a difference for this pair—often the related difference *one holds words, the other holds fruit*. In contrast, for *watch–banana split*, the modal commonality was *both are long and thin*. Only 2 of the 4 participants who found this commonality listed a difference, and the differences listed were nonalignable (e.g., *only one is food, only one goes on your wrist*).

Although the distinction between relational and attributional similarity appears subtle, it is psychologically important (Bassok, Wu, & Olseth, 1995; Gentner, 1988; Gentner & Clement, 1988; Goldstone et al., 1991; A. B. Markman & Gentner, 1993b, 1996; Medin et al., 1990; Ross, 1989). For example, when the participants interpreted metaphors and rated their aptness, their aptness ratings increased with the relationality of the interpretations, but were independent or negatively correlated with their attributionality (Gentner, 1988; Gentner & Clement, 1988). There is other evidence for a psychological distinction between relational and attributional similarity. In two-alternative similarity choice tasks, Medin et al.

(1990) found differential weighting: Relational commonalities were weighted strongly in similarity judgments and attributional commonalities in difference judgments. Goldstone et al. (1991) found that people's similarity matches were biased toward whichever similarity type—*attributional or relational*—was most strongly represented in the two alternatives. Gentner, Rattermann, and Forbus (1993) found that ratings of inferential soundness and of similarity were higher for story pairs that shared higher order relational structure, whereas the probability of being reminded of one story by another depended on the level of attributional similarity. Relational similarity appears to become more salient and preferred with age and experience (Gentner & Rattermann, 1991; Gentner & Toupin, 1986; Halford, 1987; Rattermann & Gentner, 1998). These findings suggest that relational commonalities differ from attributional commonalities and, at least for adults, contribute more strongly to a subjective sense of aptness and similarity.

If indeed relational similarity matches are larger and deeper, structure-mapping theory predicts that the benefits of prior alignment experience should be most pronounced for relational similarity pairs. We tested this prediction in Experiment 3B by contrasting relational and attributional pairs. For this experiment, we contrasted low-similarity pairs that were specifically relational or attributional, along with pairs of overall high similarity. The relational pairs were intended to elicit a single common relation (e.g., *boa constrictor* and *girdle* both constrict something), and the attributional pairs were intended to elicit a single common attribute (e.g., *sandpaper* and *stubble* are both rough). The word pairs were created according to the experimenters' intuitions and their relationality and attributionality were later confirmed by expert raters, as described below (Experiment 3A).

For Experiment 3B, we predicted, as before, that more differences would be produced for high-similarity pairs than for low-similarity pairs. A further prediction was that relational pairs would yield more differences than attributional pairs. More importantly, the difference facilitation for old pairs over new should be greater for relational pairs than for attributional pairs. Lastly, if the relational advantage results from larger common systems, the differences produced for relational pairs should be more meaningfully related to the commonalities than should those produced for attributional pairs (Experiment 3C).

### EXPERIMENT 3A

We first needed to ensure that we had a contrasting set of pairs capturing attributional and relational similarity, respectively. To do this, we began with an initial set of 40 relational and 40 attributional pairs and winnowed them down to 18 of each type by the following method. We first obtained commonalities from 64 participants for the initial set of 80 pairs. Each participant listed commonalities for high overall similarity pairs (included as literal similarity anchors) and either relational or attributional pairs (this varied between subjects). Then, we asked two

**Table 4**  
**Relational and Attributional Similarity Pairs**  
**Used in Experiment 3**

Relational Similarity Pairs		Attributional Similarity Pairs	
Fog	Mask	Railroad tracks	Ladder
Repair shop	Hospital	Pancake	Plains
Vampire	Leech	Hair	Spaghetti
Outlet	Electric eel	Lemon	Sun
Venus flytrap	Spider web	Pickle	Astro turf
Telescope	Radar	Sandpaper	Stubble
Earthquake	Egg beater	Cloud	Fleece
Bleach	Confession	Half-moon	Scythe
Weeds	Graffiti	Football	Egg
Exam	Filter	Piano keyboard	Zebra
Ladder	High-heeled shoes	Barbershop pole	Tiger
Blinds	Sunglasses	Eraser	Tires
Road beam	Spinal column	Satellite dish	Birdbath
Traffic ticket	Spanking	Curly hair	A slinky
Elevators	Arteries	Silk	Sheet metal
Lungs	Accordion	Doorknob	Mushroom
Boa constrictor	Girdle	Skunk	Garlic
Helium balloon	Bread dough	Flea	Atom

trained undergraduate raters, naive to the conditions and hypotheses of the experiments, to judge the attributionality and relationality of the commonalities produced for a random half of the stimuli using the method developed by Gentner and Clement (1988). (Commonalities for the high-similarity pairs were not rated.) The raters were told that *attributionality* refers to the degree to which a predicate describes objects in and of themselves (e.g., *both are square*), whereas *relationality* is the degree to which a predicate describes relations between objects or between relations (e.g., *both allow you to cross water*). They were trained with 10 practice examples before moving on to the actual ratings task. For each distinct commonality, the raters judged the degree of relationality and attributionality on separate 1–5 scales. The raters were not given the stimulus items themselves. The judges' ratings were within 1 rating point of each other 88% of the time, and discrepancies were resolved by discussion.

The relationality ratings for the two new similarity types were as follows: relational pairs,  $M = 3.88$ ; attributional pairs,  $M = 1.76$  (out of 5). The attributionality ratings were, for relational pairs,  $M = 3.38$ , and for attributional pairs,  $M = 4.50$ . Paired sample  $t$  tests revealed that these ratings differed across these two types of similarity pairs: for relationality [ $t(19) = 15.79, p < .001$ ] and for attributionality [ $t(19) = 8.32, p < .001$ ]. Thus, the materials satisfied the task requirements: The commonalities generated for the attributional pairs earned high scores for attributionality and low scores for relationality, and the reverse was true for the relational pairs; they were rated high in relationality and only moderate in attributionality. The attributional and relational pairs used in this experiment are shown in Table 4.

### EXPERIMENT 3B

#### Method

**Participants.** Thirty-two Northwestern University undergraduates participated. None had participated in the earlier experiments.

**Materials, Design, and Procedure.** The materials were relational, attributional, and overall-similarity word pairs. The overall-similarity pairs were the high-similarity pairs used in Experiment 1 and were included as literal similarity anchors to help maintain natural similarity processing. The participants were asked to list a single commonality for 27 word pairs (9 each of overall, relational, and attributional similarity) and then to complete a speeded difference task on 54 pairs (18 of each similarity type) that were half old and half new. The two stimulus subsets making up the commonality task were presented in four random orders. The full stimulus set, used in the difference task, was seen in four random orders and their reverse orders. The design was thus similarity type (overall, relational, or attributional, within-subjects)  $\times$  experience level (old or new, within-subjects)  $\times$  stimulus subset (between-subjects).

#### Results and Discussion

Similarity type affected the number of differences generated (Figure 2A). The number of differences listed were as follows: overall similarity pairs,  $M_{old} = 5.5, M_{new} = 5.4$ ; attributional pairs,  $M_{old} = 5.6, M_{new} = 4.8$ ; relational pairs,  $M_{old} = 5.1, M_{new} = 4.2$  [ $F_s(2,62) = 4.85, MS_e = 2.48, p < .05$ ]. Overall, difference listing was facilitated for old pairs ( $M = 5.4$ ) relative to new pairs ( $M = 4.8$ ) [ $F(1,31) = 9.55, MS_e = 1.77, p < .01$ ], as expected. This difference facilitation appears larger for the relational pairs, but was reliable for both the relational and the attributional pairs: for relational pairs [ $t(31) = 2.93, p < .01$ ]; for attributional pairs [ $t(31) = 2.82, p < .01$ ] (both by Bonferroni adjustment). No interactions were found. An item analysis also showed a main effect for similarity type [ $F_i(2,48) = 4.33, MS_e = 4.94, p < .05$ ] and experience [ $F_i(1,48) = 12.06, MS_e = 2.50, p < .01$ ], as well as an interaction between experience and stimulus set [ $F_i(1,48) = 10.42, MS_e = 2.50, p < .01$ ]. Despite showing considerable difference facilitation (for old over new), the relational pairs did not appear to elicit as many overall differences as did attributional pairs. Such results appear to contradict the structural alignment prediction that larger alignments (as with the relational pairs) would lead to greater ease in producing related differences.

However, an examination of the kinds of differences produced suggested that the differences listed for the attributional pairs were of a rather "formulaic" nature. First, we noticed that for attributional pairs, the participants often used a small set of generic differences (e.g., *one is bigger*) across a number of word pairs. The generic dimensions most often used were size, animacy, natural vs. man-made, and edibility. The responses of one participant, who relied heavily on this strategy, are shown in Table 5. This response pattern had occurred only rarely in previous experiments. Such repeated use of the same dimensions might suggest a rather perfunctory approach to generating differences.

Another formulaic pattern frequently used with attributional pairs was simple category labeling, in which the participants who were asked to give a difference between two items simply supplied their category names. For example, given the word pair *football-egg*, a participant might write *sport-food*. (See Table 6 for other examples from a single participant.) The participants appear to be saying, "these two objects are different because they are different kinds of things"—again, a rather lazy strategy.

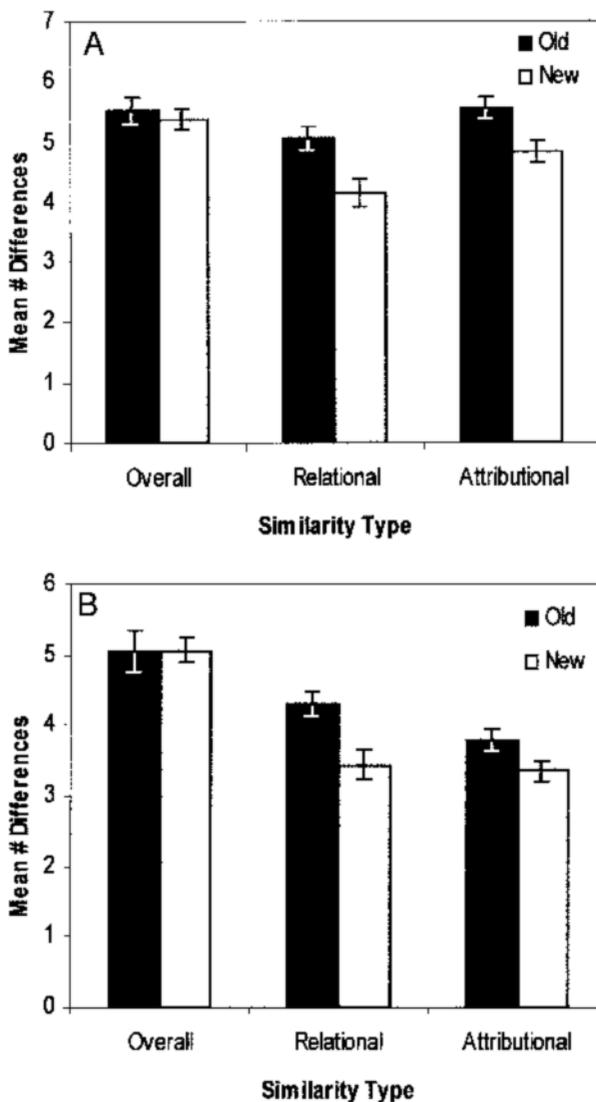


Figure 2. (A) Results of Experiment 3B: Mean differences  $\times$  similarity type and experience. (B) Results of Experiment 3B: Mean differences  $\times$  similarity type and experience for specific (nonformulaic) differences only.

In contrast, the differences listed for relational pairs appeared related to specific functional aspects of particular pairs. The repeated use of simple category labels and a small number of generic differences (together, “formulaic” responses) can be seen as an efficient way of dealing with pairs, such as attributional pairs, for which generating differences is inherently very difficult, because the alignment is poor. That is, we suspected that this behavior represents a compensatory strategy for low alignability. To determine whether this was indeed the case, we carried out further ratings and analyses.

**Formulaic responses.** Our first goal was to verify whether formulaic strategies were indeed associated more with attributional pairs than with relational pairs. We scored a difference as formulaic if it was given as a participant’s second or subsequent use of a particular generic

difference (e.g., size, animacy) or his/her second or subsequent use of category labels (e.g., one is an  $x$ , one is a  $y$ ). (The first use was not counted as formulaic. For consistency, the “first” use was assumed to be the first such difference found scanning down the page from the top left, even though the participants were free to skip around on the page.) The number of formulaic differences was recorded for each pair type.

Table 7 displays the percentage of differences generated for each pair type that was formulaic, averaged across experience level. The percentage of formulaic differences increased as alignability decreased: for attributional pairs, 16.5%; for relational pairs, 8.3%; and for overall-similarity pairs, 2.3% [ $\chi^2(2) = 11.60, p < .05$ ]. The table also shows the percentage of participants who used four or more of any kind of generic or category labeling within a single pair type. One can see that the participants were more likely to use formulaic differences repeatedly for attributional pairs than for any other pair type.

We return now to our original predictions for Experiment 3. We proposed that relational pairs should elicit more alignable differences and should show a greater old–new advantage in alignable differences, because of the greater size and depth of their common systems (relative to attributional pairs). This prediction rests on the assumption that the differences are suggested by the alignment. However, perhaps the attributional pairs were so meager in their aligned structures that the participants simply bypassed the usual difference-generation process by using formulaic strategies. Thus, the large number of formulaic responses for the attributional pairs may have obscured an alignment effect. Therefore, we reanalyzed the results of Experiment 3 by looking at only specific (nonformulaic) differences, as are shown in Figure 2B. As before, a difference was considered formulaic—and thus was dropped from the present analysis—if it was given as a participant’s second or subsequent use of a simple category label (as in the pair *a fruit* vs. *a star*) or of the same generic dimension (e.g., size).

As predicted, the number of nonformulaic differences listed was as follows: overall-similarity pairs,  $M_{old} = 5.1, M_{new} = 5.1$ ; relational pairs,  $M_{old} = 4.3, M_{new} = 3.4$ ; and attributional pairs,  $M_{old} = 3.8, M_{new} = 3.3$  [ $F(2,60) = 15.42, MS_e = 2.60, p < .001$ ]. Across all conditions, dif-

Table 5  
Example Use of Generic Differences  
in Experiment 3B (Subject 27)

doorknob/mushroom:	you can eat one
eraser/tires:	size
barbershop pole/tiger:	one is alive
yacht/sailboat:	size
piano keyboard/zebra:	living/not living
weeds/graffiti:	living/not living
football/egg:	one is edible
earthquake/eggbeater:	natural/man-made
pickle/astroturf:	one was alive
lemon/sun:	eat one
venus flytrap/spider web:	one is alive
outlet/electric eel:	natural/man-made
store/boutique:	size

**Table 6**  
**Example Use of Simple Category Labeling**  
**in Experiment 3B (Subject 32)**

lemon/sun:	fruit/star
pickle/astroturf:	vegetable/fake grass
earthquake/egg beater:	disaster/appliance
football/egg:	sport/food
piano keyboard/zebra:	instrument/animal
lungs/accordion:	organ/instrument
silk/sheet metal:	cloth/metal
boa constrictor/girdle:	snake/cloth
skunk/garlic:	animal/spice
flea/atom:	bug/scientific object

ference listing was again facilitated for old pairs ( $M = 4.4$ ) relative to new pairs ( $M = 3.9$ ) [ $F(1,30) = 4.37$ ,  $MS_e = 2.11$ ,  $p < .05$ ]. The most important finding, however, was that the difference facilitation was reliable for the relational pairs [ $t(31) = 2.4$ ,  $p < .05$ ] but not for the attributional pairs [ $t(31) = 1.29$ ,  $p = .2$ , n.s.] (both by Bonferroni adjustment). Interestingly, the overall-similarity pairs also failed to show an experience effect; the participants produced large numbers of differences (whether scored as total differences or as nonformulaic differences) for both old and new pairs. As we speculated earlier, high-similarity pairs may already have differences stored, or they may be so readily alignable that the prior similarity task is superfluous. Both the large number of differences for high-similarity pairs and the greater facilitation effect for relational pairs than for attributional pairs are consistent with the claim that the extent of the common system influences the degree of difference facilitation.

**Differences in relation to commonalities.** Structural alignment theory makes predictions not only about the number of differences, but also about the kind of differences that should be generated. Specifically, the theory predicts that the comparison process should yield differences that are related to, or derived from, the commonalities (A. B. Markman & Gentner, 1993a). As noted above, prior studies have found not only that more differences are produced for high-similarity pairs than for low-similarity pairs (Gentner & Markman, 1994), but that this high-similarity advantage resides in *alignable differences* (typically stated as different values on a common predicate, as in *bicycles have two wheels, cars have four*) rather than in *nonalignable differences* (typically stated by asserting a fact for one item and denying it for the other, as in *bicycles have wheels, apples don't*).

As noted earlier, A. B. Markman and Gentner (1993a) verified the greater connectivity of alignable differences by asking independent judges to sort the commonalities and differences generated by previous subjects. The results showed many conceptual relationships between alignable differences and commonalities, but few relationships between nonalignable differences and commonalities. We tested the corresponding prediction in the present experiment. Because the relationally similar pairs should have large systems of commonalities (relative to the attributional pairs), the structural alignment process should give

rise to more related differences for these pairs. Thus, the differences listed for relational pairs are more likely to be conceptually related to the commonalities listed for the same pairs than are the differences listed for attributional pairs. This prediction was tested in Experiment 3C.

## EXPERIMENT 3C

### Method

**Participants.** The participants were 8 undergraduates from Northwestern University, who participated in partial fulfillment of a course requirement. None had participated in the earlier experiments.

**Materials.** The materials consisted of participant responses from one randomly chosen stimulus set from Experiment 3B (i.e., responses from 16 participants for nine relational and nine attributional pairs). The experimenter classified responses for each pair as *consensus* commonalities or differences if they had been produced by 2 or more participants; all other responses were discarded. Of the 288 individual *commonality* listings (1 from each of 18 pairs for each of the 16 participants), 274 responses were retained as consensus responses, and 14 were discarded. Every pair elicited slight variations of a single commonality across all participants (e.g., for *Venus flytrap-spider*, all responses were slight variations of the same idea: *both trap things, both catch things, both trap prey*, etc.). This tendency to settle on one common system is consistent with findings that suggest that as the winning alignment emerges, predicates not consistent with this alignment are suppressed (Gernsbacher, Keysar, & Robertson, 1995). A few pairs elicited two commonalities (e.g., for *satellite dish-birdbath*, some participants listed variations of *concave, dish-shaped*, whereas others listed variations of *both are outside*). Thus, the 274 consensus responses were collapsed into one or two nonredundant consensus commonalities per pair (23 across the 18 pairs).

There were fewer differences listed overall, because the participants had been given limited time to perform the difference task. Also, because the pairs were of low similarity, the differences elicited were more variable than were the commonalities. Of the 161 individual difference listings, 123 were retained as consensus differences, and 38 were discarded. The 123 differences were collapsed into one to three nonredundant differences per pair (33 across the 18 pairs). For example, for *earthquake-eggbeater*, the consensus differences were *one more dangerous, one more subtle and one beats eggs, the other shakes the earth*. For *flea-atom*, the consensus differences were *one is an insect, one is alive, and one is larger than the other*.

The consensus commonality or commonalities for each pair were written on a single index card, for a total of 9 relational pair commonality cards and 9 attributional pair commonality cards. Each consensus difference was written on a separate card, for a total of 12 relational pair difference cards and 21 attributional pair difference cards. The experimenter kept track of which pairs the com-

**Table 7**  
**Frequency of Use of Formulaic Strategies in Experiment 3B**

Similarity Type	Formulaic Differences		% Subjects
	% Labeling	% Generics	
Overall	1.5	3.0	0
Relational	12.5	4.0	4.7
Attributional	16.5	16.4	14.1

Note—Similarity types are averaged over experience level. % Labeling and % Generics are the percentages of total differences listed that were formulaic. % Subjects is the percentage of participants out of 32 who used four or more of any particular generic and/or category labeling response for the given pair category.

monalities and differences came from, but the pairs themselves did not appear on any of the sorting materials.

**Design and Procedure.** Each participant was run individually in two phases. In both phases, the procedure was the same. The consensus commonalities were spread out on the table in front of the participant, and a stack of consensus difference cards was given to him/her. The participant was told that each card on the table represented a commonality or commonalities that previous participants had listed for a particular pair of objects. (The object names were not given.) The participant was also told that the cards in the stack he/she was holding were differences listed for the same set of objects. The participant was instructed to judge whether or not each difference was "conceptually related" to any of the commonalities. They were to place each difference card on the appropriate commonality card to indicate relatedness or to the side to indicate a lack of relatedness to any commonality. The instructions made clear that a given commonality card might have several, one, or no differences related to it, and that there were no "right or wrong" answers. One (nonscored) practice trial was run with different materials to ensure that each participant understood and was comfortable with the procedure.

The sorting task was carried out twice: once for the set of relational pairs and once for the set of attributional pairs. (The two similarity types were separated to avoid contrast effects.) Half of the participants sorted the relational responses first, and half sorted the attributional responses first. The sorting participants were not told that the two sets differed in any way. The dependent measure was simply the proportion of times that a difference was matched with a commonality from the same word pair.

## Results

As predicted, this measure differed between relational and attributional pairs. The differences for a given pair were matched to the commonalities for that same pair 87.5% of the time for relational pairs, but only 42.9% of the time for attributional pairs [ $t(7) = 8.41, p < .01$ ]. Only the differences produced for relational pairs show specificity of connection to common structure.

After the sorting task, the participants were asked whether they thought that the two groups of materials differed in any way. Sorting the attributional responses was universally considered to be more difficult than sorting the relational responses. Some of the participants could not say why. Others stated that the commonalities or differences were more specific for the relational task. Three participants remarked that one "could have put the differences anywhere" on the attributional task. Another stated that the attributional differences were "too general; there was not enough information to choose one commonality as more related than another."

Thus, the relational pairs showed clear difference facilitation and elicited differences that were specifically related to the commonalities produced for the same pairs. Neither of these held for the attributional pairs. As predicted, deep alignments facilitated the noticing of differences that were conceptually connected to the common structure.

## GENERAL DISCUSSION

In these studies, we tested three predictions of structural alignment theory with respect to differences. The first prediction was that aligning two concept represen-

tations should facilitate finding differences between them. Second, this facilitation should be specifically related to the structural alignment process and should not simply result from joint activation of the word pairs. Third, the degree of difference facilitation should reflect the depth of the common system. Fourth, pairs with deep and rich alignments should elicit not only more differences, but more differences specifically related to the commonalities, than pairs with more sparse alignments. The present experiments provide support for these hypotheses.

In Experiment 1, participants were able to list more differences for pairs they had recently aligned than for pairs they had not, suggesting that structural alignment facilitates noticing differences. Consistent with prior research, there was also an advantage for high-similarity pairs over low-similarity pairs. Experiment 2 supported our second prediction by demonstrating that mere co-processing of the pairs was not sufficient to produce this effect. Only the participants in the comparison group listed more differences for old than for new pairs. Although the participants in the thematic relation group had processed the old pairs, they received no difference-listing advantage from such nonalignment experience, but rather showed diminished performance. Concerning the third hypothesis, when only the specific, nonformulaic responses in Experiment 3 were analyzed, more differences were listed for relational pairs than for attributional pairs, and the difference facilitation was reliable only for the relational pairs. The fourth hypothesis was also borne out: Relational pairs, with deeper common systems, inspired differences more specific to particular comparisons than did attributional pairs, with more shallow common systems. These results, and the high-similarity advantage in Experiment 1, suggest that success and speed in finding differences is related to the size of the common system extracted. These findings bear out the claim that the structural alignment process highlights commonalities and the differences related to those commonalities.

The present results extend previous work that has explored the relationship between commonalities and related differences. A number of studies (A. B. Markman & Gentner, 1993a, 1996; A. B. Markman & Wisniewski, 1997) have demonstrated that, somewhat paradoxically, people can generate a greater number of alignable differences for high-similarity than for low-similarity word pairs or pictures. Strikingly, people list a greater number of differences (of any sort) under time pressure for high-similarity than for low-similarity word pairs (Gentner & Markman, 1994). What remained to be demonstrated, however, was that the structural alignment process per se is responsible for the observed psychological availability of differences. The present experiments demonstrated this link directly, in that (1) word pairs that had been previously aligned were those that were easiest to list differences for, and (2) comparison, rather than mere coactivation, specifically led to this difference facilitation.

### Thematic Relatedness Versus Similarity

The results of Experiment 2—that differences are easily derived from commonalities but not from thematic relations—bear on the larger issue of the relation between similarity and thematic relatedness. A basic assumption in cognitive theorizing is that thematic associations and similarity relationships are distinct—that is, the similarity relation between *milk* and *cream* is of a different kind and presumably results from a different process than the thematic relation between *milk* and *cow*. Yet some recent theories suggest a strong link between similarity and thematic association (Bassok & Medin, 1997; Gentner & Brem, 1998; Wisniewski & Bassok, 1996). For example, A. B. Markman and Gentner (1996) noted that in commonality-listing tasks, people sometimes list thematic relations for low-similarity pairs (e.g., that *kitten* and *ball-bearing* have the commonality that a kitten could play with a ball-bearing). Thematic relations can also influence similarity judgments. Wisniewski and Bassok compared the similarity ratings assigned to similar pairs (*milk–lemonade*), thematic pairs (*milk–cow*), and pairs sharing both similarity and thematic relationships (*milk–coffee*). They found that thematic relationships significantly increased similarity ratings (e.g., *milk* and *coffee* were rated as more similar than *milk* and *lemonade*). Gentner and Brem (1998) gave participants similarity triads in which one alternative (e.g., *shovel*) was thematically related to the standard (e.g., *snow*) and the other (e.g., *rain*) was similar to the standard. The participants sometimes chose the thematically related alternative, even though they were explicitly instructed to choose on the basis of similarity. Further, the response times for correct performance were slowed by the presence of a thematic competitor. Bassok and Medin (1997) used sentence pairs to show that alignability, rather than undifferentiated similarity, may determine when thematic relations intrude on similarity judgments. When the sentence pairs shared a common verb, the participants justified their similarity responses on the basis of the structural alignment of relational structure. But when the sentence pair shared only a common noun, as in *The carpenter fixed the chair/The carpenter sat on the chair*, the justifications often were based on thematic links (e.g., that the carpenter could sit on a chair to test whether it was fixed). These findings, taken together, show that when alignment is difficult, similarity ratings can be influenced by thematic relations.

These findings have led to questions about the nature of similarity and thematic relatedness. Bassok and Medin (1997) raised the possibility that the term *similarity* should encompass a general cognitive sensation of relatedness that includes both commonalities and thematic associations. Such an account is consistent with Sloman's (1996) two-system proposal, in which similarity and associative relationships belong to a single associative system, which is separate from, and sometimes competes with, a rule-based system. In contrast, Gentner and Brem (1998, 2001) proposed that similarity and thematic relatedness result

from two separate processes—comparison and retrieval of associations, respectively—whose output is sometimes hard to differentiate through direct introspection, particularly with low-similarity pairs whose alignment is not particularly salient. (An additional complexity is that a similarity relation can be cached as a stored association if computed sufficiently often.) In short, people experiencing a strong sense of relatedness may not always be able to explicitly label whether its source is alignment or retrieval. Gentner and Brem (1998, 2001) used a word extension task to show that people's selectivity improved when the need for direct introspection was removed. It is well established that word extensions are based on likeness, not thematic associations, even for preschool children (E. M. Markman & Hutchinson, 1984; Waxman & Gelman, 1986). Applying this technique to adults, Gentner and Brem found that adults in a word extension task ("In a foreign country, the people call spoons *blinks*. Which of these is also a *blink*?") almost always chose the similar alternative over the thematic alternative (2% thematic errors). Consistent with the confusability hypothesis, the same materials in a direct similarity task (described above) garnered many more thematic errors (15%–30%, depending on deadline).

The results of the present Experiment 2 argue strongly that comparison and thematic association are separate processes. People were more fluent at producing differences after stating commonalities for the same pair, and less fluent after producing a thematic relation. Discovering a thematic association actually retards the process of extracting differences relative to no prior task. The difference-listing method bypasses the need for people to make an explicit introspective distinction between two kinds of relatedness; instead, we traced the effects of the process on a subsequent task. The fact that producing commonalities and producing thematic relations have opposite effects on generating differences is strong evidence that they are distinct processes.

### Conclusions

Research has shown that differences are more easily generated for high-similarity than for low-similarity comparisons. The present experiments extend and clarify these findings by showing that the structural alignment process per se facilitates finding differences. Mere prior coprocessing of the concepts is not sufficient for the effect. Further, pairs with deeper common systems elicit a larger number of specific alignable differences than do pairs with more shallow alignments. In sum, the alignment process illuminates both common structure and differences related to that common structure.

These results have broader implications. The nonequivalence of commonalities and differences challenges mental space representations, in which difference and similarity are simply inverses. The close relationship between commonalities and differences challenges independent-feature models of representation, which have no way to express this connectivity. Of course, different modeling

choices are useful for different aspects of mental phenomena. The present results show that even so basic a cognitive process as noticing a difference makes use of structured representations and structural alignment processes.

## REFERENCES

- BASSOK, M., & MEDIN, D. L. (1997). Birds of a feather flock together: Similarity judgments with semantically rich stimuli. *Journal of Memory & Language*, **36**, 331-336.
- BASSOK, M., WU, L.-L., & OLSETH, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem solving transfer. *Memory & Cognition*, **23**, 354-367.
- CLEMENT, C. A., & GENTNER, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science*, **15**, 89-132.
- FALKENHAINER, B., FORBUS, K. D., & GENTNER, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, **41**, 1-63.
- GENTNER, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, **7**, 155-170.
- GENTNER, D. (1988). Metaphor as structure-mapping: The relational shift. *Child Development*, **59**, 47-59.
- GENTNER, D., & BREM, S. (1998). Is snow really like a shovel? Distinguishing similarity from thematic relatedness. *Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society* (pp. 179-184). Mahwah, NJ: Erlbaum.
- GENTNER, D., & BREM, S. (2001). *Alchemy in current cognition: Distinguishing similarity from thematic relatedness*. Manuscript in preparation.
- GENTNER, D., & CLEMENT, C. (1988). Evidence for relational selectivity in the interpretation of analogy and metaphor. *Psychology of Learning & Motivation*, **22**, 307-358.
- GENTNER, D., & MARKMAN, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science*, **5**, 152-158.
- GENTNER, D., & MARKMAN, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, **52**, 45-56.
- GENTNER, D., & RATTERMANN, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Brynes (Eds.), *Perspective on thought and language: Interrelations in development* (pp. 225-277). London: Cambridge University Press.
- GENTNER, D., RATTERMANN, M. J., & FORBUS, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, **25**, 524-575.
- GENTNER, D., & TOUPIN, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, **10**, 277-300.
- GERNSBACHER, M. A., KEYSAR, B., & ROBERTSON, R. R. (1995). *The role of suppression in metaphor interpretation*. Paper presented at the Thirty-Sixth Annual Meeting of the Psychonomic Society, Los Angeles.
- GOLDSTONE, R. L., MEDIN, D. L., & GENTNER, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology*, **23**, 222-262.
- HALFORD, G. S. (1987). A structure-mapping approach to cognitive development. The neo-Piagetian theories of cognitive development: Toward an interpretation. *International Journal of Psychology*, **22**, 609-642.
- JAMES, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.
- MARKMAN, A. B., & GENTNER, D. (1993a). Splitting the differences: A structural alignment view of similarity. *Journal of Memory & Language*, **32**, 517-535.
- MARKMAN, A. B., & GENTNER, D. (1993b). Structural alignment during similarity comparisons. *Cognitive Psychology*, **25**, 431-467.
- MARKMAN, A. B., & GENTNER, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition*, **24**, 235-249.
- MARKMAN, A. B., & GENTNER, D. (1997). The effects of alignability on memory storage. *Psychological Science*, **8**, 363-367.
- MARKMAN, A. B., & WISNIEWSKI, E. J. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 54-70.
- MARKMAN, E. M., & HUTCHINSON, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, **16**, 1-27.
- MEDIN, D. L., GOLDSTONE, R. L., & GENTNER, D. (1990). Similarity involving attributes and relations: Judgments of similarity and differences are not inverses. *Psychological Science*, **1**, 64-69.
- MEDIN, D. L., GOLDSTONE, R. L., & GENTNER, D. (1993). Respects for similarity. *Psychological Review*, **100**, 254-278.
- NOSOFSKY, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 87-108.
- RATTERMANN, M. J., & GENTNER, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, **13**, 453-478.
- ROSS, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier examples. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 456-468.
- SHEPARD, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, **39**, 373-421.
- SHOBEN, E. J. (1983). Applications of multidimensional scaling in cognitive psychology. *Applied Psychological Measurement*, **7**, 473-490.
- SLOMAN, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, **119**, 3-22.
- TVERSKY, A. (1977). Features of similarity. *Psychological Review*, **84**, 327-352.
- WAXMAN, S. R., & GELMAN, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development*, **1**, 139-156.
- WISNIEWSKI, E. J., & BASSOK, M. (1996). On putting milk in coffee: The effect of thematic relations on similarity judgments. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 464-468). Mahwah, NJ: Erlbaum.

## NOTES

1. This generalization fails, of course, at the point of identity (no differences), leading to the intriguing question of the exact nature of the limit sequence.

2. Due to experimenter error, a few of the particular words used in certain relational pairs were also used in a different form in certain attributional pairs (e.g., *piano*, *piano keyboard*). However, this affected only a few of the stimulus sets, and any impact was likely to have been neutralized by the randomizing procedures.

**APPENDIX**  
**High- and Low-Similarity Word Pairs From Experiment 1**

High Similarity		Low Similarity	
Staple	Paper clip	Store	Hang glider
Kitten	Cat	Yacht	Mop
Chemistry	Biology	Stairs	Boutique
Shoe	Sandal	Police car	Painting
Light bulb	Candle	Kite	Burger King
Magazine	Newspaper	Rocket	Motel
Bicycle	Tricycle	Casino	Clock
Telephone	CB radio	Bed	Missile
Piano	Organ	Watch	Banana split
Freezer	Refrigerator	Hammock	Horse track
Computer	Typewriter	Hotel	Hockey
Dumpster	Garbage can	Chair	Tape deck
Air conditioner	Furnace	Ice cream sundae	Escalator
Lake	Ocean	Football	Stool
Phone book	Dictionary	Sculpture	Navy
Bowl	Mug	Calculator	Lounge chair
Hammer	Mallet	McDonald's	Couch
Sponge	Towel	Broom	Ambulance
Diamond	Ruby	Army	Abacus
Chair	Stool	Shoe	Paper clip
Watch	Clock	Bowl	Dictionary
VCR	Tape deck	Freezer	Typewriter
Casino	Horse track	Piano	Tricycle
Sculpture	Painting	Computer	Sandal
McDonald's	Burger King	Light bulb	Cat
Yacht	Sailboat	Air conditioner	Mallet
Ice cream sundae	Banana split	Dumpster	CB radio
Police car	Ambulance	Telephone	Ruby
Hammock	Lounge chair	Chemistry	Candle
Bed	Couch	Magazine	Mug
Store	Boutique	Kitten	Ocean
Kite	Hang glider	Hammer	Biology
Football	Hockey	Staple	Furnace
Calculator	Abacus	Bicycle	Refrigerator
Broom	Mop	Sponge	Organ
Hotel	Motel	Phone book	Stereo speaker
Stairs	Escalator	Microphone	Garbage can
Army	Navy	Lake	Newspaper
Rocket	Missile	Diamond	Towel

(Manuscript received May 5, 1999;  
revision accepted for publication July 3, 2000.)