Learning by Analogical Bootstrapping

Kenneth J. Kurtz, Chun-Hui Miao, and Dedre Gentner

Department of Psychology
Northwestern University

Analogies are typically drawn from a well-understood situation to a situation that is poorly understood. In this research, we investigate a different route to analogical insight. We suggest that mutual alignment—that is, comparison between 2 partially understood situations—can act to promote comprehension and abstraction. We presented participants with 2 analogous scenarios depicting heat flow. They were given processing tasks that varied in the degree to which comparison was required. We then measured insight into the common structure in 3 ways. Participants were asked to (a) specify differences between the 2 pictured scenarios, (b) write scenario descriptions, and (c) rate the similarity of the 2 scenarios. The results show that carrying out comparison promotes greater insight into the common causal structure, but only when the comparison is intensive. The best results were obtained when participants were asked to jointly interpret the scenarios and to list specific correspondences. In a second experiment designed to further pinpoint the source of the comparison advantage, participants were asked to make correspondences between the elements of the 2 scenarios. These results suggest that mutual alignment is an effective means of promoting insight.

The ability to think analogically is central to human cognition. Progress in scientific and technological domains often arises from the discovery of a profound analogy such as the planetary model of atomic structure, the water flow analogy of heat transfer, and the computer metaphor of the human mind (Gentner et al., 1997; Hesse, 1966; Nersessian, 1992; Thagard, 1992a). Dunbar (1993, 1995) suggested, on the basis of his studies of the behavior of scientists in real-world microbiology laboratories, that analogy use is positively correlated with successful scientific investigation.

Correspondence and requests for reprints should be sent to Kenneth J. Kurtz, Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208–2710. E-mail: kjk@nwu.edu

Research on analogical processing and case-based reasoning shows the importance of analogy in learning and understanding (Holyoak & Thagard, 1995; Kolodner, 1993, 1997; Nersessian, 1992; Reed, 1987; Reeves & Weisberg, 1994; Ross, 1984, 1986; Thagard, 1992b). Students can often understand a novel situation by transferring knowledge from a well-understood situation (Pirolli & Anderson, 1985; Polya, 1945; Riesbeck & Schank, 1989). Ross (1987) found that giving learners analogical examples to illustrate a probability principle facilitated their later use of the probability formula to solve other problems. In sum, there is good evidence that analogy can serve as a powerful learning tool.

The standard approach to analogical learning involves mapping information from a base— a well-understood domain that serves as the information source—to a less familiar target. The base analog is assumed to be rich in information and to have well-understood relational structure; in contrast, the target is not well understood. In this article, we explore another kind of analogical learning, which we call mutual alignment or analogical encoding (see Ferguson, 1994; Ferguson & Forbus, 1996). In mutual alignment, the learner is simultaneously presented with two analogous situations that act symmetrically in the mapping process: Both serve as potential sources and recipients of information. We suggest that this kind of comparison between two partially understood situations can lead to noticing parallel structure and developing a deeper understanding of both situations.

There are several reasons for exploring mutual alignment as an alternative to the standard base-to-target mapping paradigm. In the traditional framework, retrieval is typically required; examples are presented in isolation and it is assumed that learners will draw on appropriate prior examples when they are given a new target problem. On this account, analogical learning involves being reminded of prior problems. However, it has been shown that learners frequently fail to transfer relevant stored knowledge (e.g., Perfetto, Bransford, & Franks, 1983; Ross, 1989; Weisberg, DiCamillo, & Phillips, 1978). Gick and Holyoak (1980) showed that participants given an insight problem typically failed to solve it, even when they had just read a story with an analogous solution. Eighty percent of participants failed to spontaneously retrieve the analogy despite being able to recognize the analogy when reminded of the prior story. Even under conditions of recent, salient exposure to an applicable base problem, spontaneous relational retrieval is unlikely.

The difficulty of transfer has been interpreted in terms of a dissociation between the kind of similarity that governs access to long-term memory and the similarity that is used in evaluating and reasoning from a present analogical match (Gentner, Rattermann, & Forbus, 1993; Holyoak & Koh, 1987; Ross, 1989). Whereas accuracy of transfer depends critically on the degree of structural match, memory retrieval is highly sensitive to surface similarity such as common object attributes and problem context. As a result, people often fail to access structurally

appropriate materials key to analogical understanding, even when such materials are present in long-term memory (Gick & Holyoak, 1980).

Apart from the difficulty of retrieval, there are further reasons why learning based on analogical processing can be hard. Whereas the main purpose of analogy is to transfer relational structure from the base to the target, the base typically also includes surface details and irrelevant content in addition to the useful relational information. Accordingly, even when a useful base is readily available, a novice learner may not be able to clearly pick out the common structure and apply it in the target domain (Gentner & Toupin, 1986; Novick & Holyoak, 1991). In some circumstances, a bridging analogy that shares similarity with both the base and the target can help the learners to understand the analogical structure (Clement, 1993), but finding an appropriate such analogy can itself be a formidable task.

We propose mutual alignment as a means of escaping such difficulties while preserving many of the benefits of analogical processing. This view emphasizes comparison as an encoding process. In mutual alignment the learner is simultaneously presented with two analogous situations that act symmetrically in the mapping process: Both can serve as sources and recipients of information. Under these conditions, comparison between two partially understood situations can lead to noticing parallel structure and developing a deeper understanding of both situations.

There is some existing evidence that comparison can facilitate learning. Gick and Holyoak (1983) found that instructing participants to describe the similarities between two analogs led to formation of abstract schemas and facilitated transfer to a further problem. Catrambone and Holyoak (1989) showed that inducing participants to compare multiple examples improved performance on long-term transfer across contexts. Cummins (1992) found that when learners were given questions that required comparison of analogous problems, they developed deeper understanding of the solution principles than did learners who received an equal number of noncomparison questions. Schwartz and Bransford (1998) conducted classroom studies to show that inducing learners to generate distinctions between contrasting cases led to the development of more differentiated knowledge structures. Loewenstein, Thompson, and Gentner (1999) found that graduate management students who compared two analogical cases were nearly three times more likely to incorporate the common strategy into a subsequent negotiation task than were students given the same cases separately. These and related findings suggest that analogical encoding can invite abstraction and promote transfer (see also Gentner & Namy, 1999; Thompson, Gentner, & Loewenstein, 2000; Vanderstoep & Seifert, 1993).

One important factor in analogical encoding may be the depth of the comparison process. Merely reading or receiving multiple cases during the same session is often not sufficient to produce comparison-based effects (Loewenstein et al., 1999). For example, Catrambone and Holyoak (1989) found transfer benefits only

when participants were given an extensive set of directed comparison instructions; merely giving them multiple source analogs without explicit instructions to compare was insufficient to support transfer. For this reason, researchers have sought ways to promote active comparison, especially early in learning. Kolodner (1997) suggested that using software tools to line up cases or examples next to each other to reveal their correspondences may facilitate the use of analogous cases.

In this research, we investigate mutual alignment—whereby partially understood examples are compared to engender deeper comprehension—as well as the type of comparison context needed to foster learning in this manner. Novice learners often have mental representations that are weakly structured (lacking causal relations and abstractions, such as understanding the concept of rain only in terms of "getting wet") and predominantly focused on surface features, in contrast to the well-structured representations of domain experts (Chi, Feltovich, & Glaser, 1981). Mutual alignment is proposed as an online means of increasing sophisticated understanding and representation through a comparison process according to which key components of meaning are made salient, roles are made explicit, and relations are made evident.

Mutual alignment has some advantages over standard approaches to analogical instruction. In one standard technique, the learner is given an analogy from a well-structured base to the desired target case. This technique requires the existence of at least one appropriately structured case that is highly familiar to the learner. Often, such a case is not available. An alternative, more bottom-up approach is case-based training, in which learners work through a series of cases design to reveal the domain principles. Such training may be ineffective if the learner fails to retrieve the appropriate prior case when encountering a new situation. (This problem may also exist for the directional mapping technique, in that learners must retrieve the base domain on subsequent encounters with the target.) In mutual alignment, the emphasis is on juxtaposing two alignable situations and inviting learners to actively identify common structure. This juxtaposition is especially helpful for novice learners who are most likely to show surface-driven retrieval (Gentner, Rattermann, & Forbus, 1993). A further advantage of mutual alignment, which we return to in the discussion, is that it's an active and engaging process for learners.

The Structural Alignment View

The theoretical foundation for mutual alignment is structure-mapping theory. It is widely agreed that analogical reasoning involves setting up correspondences between structured representational elements of two domains and transferring information guided by the common relational structure (Gentner, 1983, 1989; Holyoak, 1985; Holyoak & Thagard, 1989, 1997). According to Gentner's struc-

ture-mapping theory, the defining characteristic of analogy is that it involves the alignment of relational structure. This process is governed by the search for correspondences between connected systems of relational predicates. The computational model of structure-mapping theory, the structure-mapping engine, begins blindly with a set of local, mutually inconsistent matches and gradually coalesces these into one or a few deep, structurally consistent alignments (Falkenhainer, Forbus, & Gentner, 1989; Forbus, Gentner, & Law, 1995).

In structure-mapping theory, cases are represented in terms of sets of predicates capturing objects, attributes, and the relations that hold among them. The commonalities and differences between two situations are found by determining the maximal structurally consistent alignment between their representations. A structural alignment is characterized by one-to-one mapping (i.e., an element in one representation can correspond to at most one element in the other representation) and parallel connectivity (i.e., if elements correspond across the two representations, then the elements they govern must correspond as well). The goodness of the match is not just a matter of the number of matching elements, but also of their structure. People prefer alignments with high systematicity—that is, matches that have deeply interconnected relational systems.

Structural Alignment and Learning

Structural alignment can serve as a basis for knowledge change (see Gentner & Medina, 1998; Gentner & Wolff, 2000; Kurtz, Gentner, & Gunn, 1999). The most straightforward mechanisms follow directly from mapping: (a) highlighting is the process by which relational information comes into greater focus through comparison, and (b) inferencing is the process by which content is added to the representation of a compared case as a result of structured pattern completion. Once the mapping process is complete, the common structure of the alignment provides a basis for abstraction.

Furthermore, cases can be re-represented as part of the comparison process (Clement, Mawby, & Giles, 1994; Falkenhainer, 1990; Kotovsky & Gentner, 1996). Lacking a comparison context, learners may form domain-specific representations that are highly distinctive and idiosyncratic (Gentner & Medina, 1998). During comparison, a "drive for similarity" can lead to altering the representational elements of one or both cases toward a common internal vocabulary (Gentner & Rattermann, 1991). Re-representation consists of a change during the comparison process of the relational elements used to represent a case. For example, to presage our current research, a learner might initially represent one situation as "cooking" and another as "melting." In the course of comparison, learners may arrive at a more uniform representation of the two situations—for example, that in both cases a state change is occurring as a result of exposure to heat. In ongoing re-

search, we are evaluating the role of re-representation as a mechanism for flexible mapping based on allowing correspondences between semantically similar, but nonidentical predicates (Gentner & Kurtz, 2001).

These mechanisms underlie mutual alignment as a basis for learning. Consider a case in which a learner's initial representations emphasize surface details and are partly deficient in causal structure. During comparison, the process of establishing correspondences reveals the maximal structural match between the representations and highlights connected commonalities and differences. Re-representation during mapping can promote uniform encoding of analogs and allow semantically related but nonidentical predicates to be aligned. When a shared connected structure is more extensive in one case than the other, a candidate inference is generated. Importantly, mutual alignment reinforces and enhances encodings in both directions. Such enriched representations can then act, either individually or jointly, as a more effective cue for retrieving relevant world knowledge or additional analogs from memory (Forbus, Gentner, & Law, 1995).

Similarity and Difference in the Structural Alignment View

Structure-mapping theory has been successfully extended to the study of similarity and difference (Gentner & Markman, 1995, 1997; Goldstone, 1994a, 1994b; Goldstone & Medin, 1994; Gentner & Gunn, in press; Markman & Gentner, 1993a, 1993b, 1996; Medin, Goldstone, & Gentner, 1993). On this view, similarities and differences are highly interrelated. The output of the comparison process consists of commonalities between items and two kinds of differences: those related to the commonalities (called alignable differences) and those unrelated to the commonalities (called nonalignable differences). For example, "a cat has four legs, but a human has two" is an alignable difference; the shared dimension of number of legs is connected to the common structure shared by cats and humans. In contrast, "A cat has four legs, but people can talk" is a nonalignable difference; there is no common predicate nor a common connection to shared structure.

Highly similar pairs tend to generate more alignable differences (ADs), whereas less similar pairs tend to generate nonalignable differences (NADs). This suggests that the process of generating differences evokes a structural alignment in which common structure and differences connected to that structure become salient. Thus, paradoxically, there is a close link between commonalities and differences. People find it easier to list differences for highly similar items than for very dissimilar items (Gentner & Markman, 1994). Differences depend on commonalities. This relation takes on importance in the experimental design presented next.
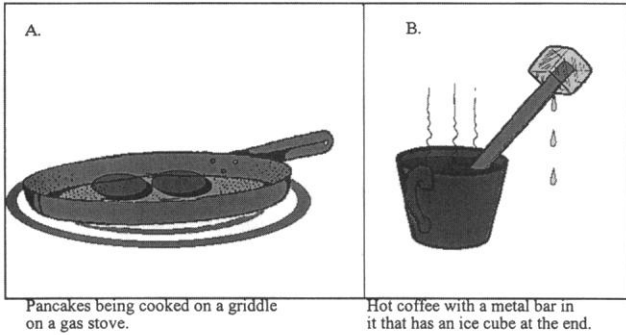
Invoking the Comparison Process

As we have noted, whether learners derive the potential advantages of mutual align-
ment may depend in part on the depth and quality of the comparison process. Many
techniques have been used to elicit comparison including physical juxtaposition of
examples (Kotovsky & Gentner, 1996), using software to generate felicitous align-
ments (Kolodner, 1997), asking for similarity ratings (Markman & Gentner, 1993b),
asking a set of directed questions (Catrambone & Holyoak, 1989), asking people to
describe commonalities (Gick & Holyoak, 1983; Loewenstein et al., 1999), and us-
ing the same name or language (symbolic juxtaposition) to refer to multiple exam-
ples (Gentner & Medina, 1998). There has been relatively little attention to calibrat-
ing effects across various techniques. Therefore, one of the goals of this research is to
further the exploration of different kinds of comparison experience.

EXPERIMENT 1

Our chief purpose in this work is to test mutual alignment as a means of facilitating
learning and understanding. In addition, we seek to specify the kinds of comparison
experiences that best promote insight. To test the influence of different types and
degrees of comparison experience, our approach is as follows:

1. Present participants with a pair of analogous situations to invite mutual
    alignment.
2. Have participants perform an orienting task designed to encourage align-
    ment during the process of interpreting and encoding the scenarios.
3. Gauge the alignment and measure resulting insight into the common causal
    structure.

The stimulus materials in this study are two pictures of analogous scenarios
shown in Figure 1: (a) pancakes in a frying pan and (b) a coffee cup with an ice
cube. Though superficially dissimilar, the two scenarios share the common theme
of heat flow. In selecting the scenarios, we made two assumptions: first, that heat
flow is the most systematic common structure underlying the two scenarios; and
second, that participants have some degree of prior understanding of the two sce-
narios and of the notion of heat flow. Because the two scenarios are quite distinct
in terms of the objects involved, the implied goal context (cooking pancakes vs.
something either to do with drinking coffee or a laboratory demonstration), and the
degree of familiarity, we expected the two cases to be at least somewhat difficult to
reconcile in alignment. More specifically, we did not expect the common structure
to be noticed effortlessly by the observer—rather, we predicted that to effectively

For each of the components of situation B write a corresponding component for situation A. (but write as many as you'd like)

Situation A.	Situation B.
1.	1. ice cube
2.	2. metal bar
3.	3. coffee
4.	4. drops of water

FIGURE 1   Two pictures of analogous scenarios.

realize and demonstrate understanding of the heat flow theme, participants must actively align the elements and their causal links.

The invitation to align the situations was varied across four conditions with different orienting tasks—that is, processing contexts within which the stimuli would be interpreted. The control condition served as a baseline and included no orienting task prior to the dependent measure of conceptual alignment. Thus, any comparison taking place would have to occur at testing. In the separate interpretation condition, participants were shown the two scenarios separately and asked to interpret them individually. In this condition, comparison is available to the participant by way of temporal juxtaposition, but it is not explicitly invited. In the joint interpretation condition, participants were shown the two scenarios side by side on the same page and explicitly instructed to compare the scenarios and describe them together. By asking participants to interpret the scenarios jointly, we expected that a learning context was provided for them to elicit the common relational structure.

Everything is in place for mutual alignment in the joint condition, but it remains a somewhat weak manipulation because the task does not specifically require a

unifying interpretation—a description could be written that addresses both scenarios, but shows minimal integration or mentions only surface features. In the joint interpretation plus correspondences condition (abbreviated Joint + Corr), participants were shown both scenarios together and asked to perform the joint interpretation task along with an additional mapping task in which specific correspondences must be assigned between the elements of the two scenarios. The joint interpretation plus correspondence task explicitly requires a comprehensive, global alignment. Accordingly, the Joint + Corr condition was expected to have the strongest effect. The order of the four conditions in terms of predicted likelihood of promoting mutual alignment was (from low to high): control, separate interpretation, joint interpretation, joint interpretation plus correspondence.

After the orienting task, we evaluated the alignment and encoding of the scenarios. An obvious measure of alignment is asking participants to list commonalities between the scenarios. We chose not to use a commonality-listing task as a dependent measure because formulating commonalities essentially involves an alignment process in its own right—consequently, effects of the orienting task manipulation might be washed out. Instead, we used a less intrusive difference listing task more likely to follow from previous processing than to initiate a new alignment. As just discussed, alignable differences tend to be produced more easily for highly similar pairs and serve as an effective measure of alignment. Differences vary in their relevance to the common structure and also in their alignability. By evaluating the kinds of differences produced we can effectively assess the degree of alignment and understanding of the common causal structure. If the participants did not find a natural way to align the stimulus, they would tend to list nonalignable or superficial differences. In contrast, successful structural alignment would enable participants to notice more alignable differences and more deep and interesting differences that are causally related to the common structure.

Raters were asked to judge the quality of each difference produced by participants in terms of depth of causal relevance to the common theme of heat flow. If the orienting task was successful in promoting mutual alignment, then participants should notice more meaningful and causally relevant differences—most of them alignable. To illustrate with an example from the experiment results, the response, "The pancakes are being heated directly from a source of heat. However, the heat in picture B is being transmitted from one substance through a medium to another substance" was rated high because of the clear account to the common structure of heat flow; whereas another response, "One is heated to be eaten, the other is a drink," received a lower rating because it is not relevant. As an additional measure of alignment, the differences were also coded based on whether the same dimension was used in both cases to delineate a different between them.

To further test the claim that mutual alignment promotes better understanding of the scenarios we collected participants' written descriptions of one of the scenarios (coffee cup with ice cube) near the end of the session. Such descriptions can

be informative about the learning taking place—Gick and Holyoak (1983) found that quality of descriptions predict subsequent performance in tests of spontaneous transfer. This description measure also allows us to investigate re-representation effects. We predict that the conditions that encourage alignment are more likely to show evidence of comparison-driven re-representation in the content of the descriptions. After completing the description task, participants were shown the pair of scenarios again and asked to rate their similarity. The judgments of similarity ought to reflect comprehension and shifts in underlying representation. We expect higher similarity ratings to result from deeper alignment, highlighting of commonalities, and increased representational uniformity.

Because alignable differences, commonalities, and similarity are positively correlated convergent measures, the results on these tasks are expected to reveal a common, consistent pattern across conditions. Compared to the difference-listing task, similarity ratings might be less informative about the specifics of alignment, but more straightforward because the results follow directly from the raw data without need for a coding process. We hope to obtain a consistent and comprehensive evaluation of mutual alignment through the use of this set of complementary measures.

METHOD

Participants

Eighty undergraduate students (half were students fulfilling a course requirement, half were recruited to participate for pay) at Northwestern University participated in the study. Twenty participants were randomly assigned to each of the four conditions.

Materials

Each participant received a packet containing the orienting task (which varied with condition) plus the set of common tasks including a difference-listing sheet, a drawing and description sheet, a similarity rating sheet, and a science background survey sheet. Two analogous scenarios depicting the common theme of heat flow (shown in Figure 1), coffee cup with ice cube and pancakes in pan, were used as the two situations to be compared. The scenarios were presented during the orienting task (except in the control condition) and appeared again on the difference-listing sheet and similarity rating sheet.

Design and Procedure

There were four different orienting tasks run in a between-subject design. Four different dependent measures were used, as shown in Table 1, that illustrate the overall

TABLE 1
Experiment 1: Design and Procedure

Conditions	Control	Separated	Joint	Joint + Correspondence
Orienting tasks	None	Separated interpretation (with pictures shown, one at a time)	Joint interpretation (with pictures shown together)	Joint interpretation then correspondence (with pictures shown together)
Common tasks (measures of alignment and learning)	Difference Listing (with pictures shown) Coffee cup drawing and description (no pictures shown) Similarity rating (with pictures shown)			
Background assessment	Science background survey			

design and ordering of the tasks. Participants were given the task sheets by the experimenter and instructed to fill them out in order. The control group began immediately with the difference-listing task that was their first exposure to the scenarios. The other groups all performed an orienting task before proceeding to the difference listing. The separate interpretation group began with an interpretation task—the scenarios were presented individually on separate pages with instructions to write a description of the scenario. The joint interpretation group started with a joint interpretation task in which both scenarios were presented on the same page. Participants were instructed to write a description as follows: "Please compare these two scenarios and think about what they have in common. Describe what is happening in these two scenarios and explain why." The Joint + Corr group started with the same joint interpretation task followed by a correspondence task in which they were shown the set of component elements of one scenario and asked to fill in the blanks with the corresponding set of elements from the other scenario.

After completing the difference-listing task, participants in all conditions were asked to draw the coffee cup scenario[1] from memory and to describe the coffee cup scenario (on the same task sheet). Participants were then asked to rate the similarity between the two scenarios on a scale from 1 (*low*) to 7 (*high*). After all of the dependent measures were collected, participants rated their own science backgrounds from 1 (*low*) to 5 (*high*). In keeping with our assumption that participants had some basic understanding of heat transfer, participants who responded in the science background self-rating were removed from the analysis and replaced. In all conditions, the phases of the task were completed without interruption in a single experimental session lasting approximately 10 min.

---

[1]This requirement was used to check whether they carefully studied the scenarios.

Results and Discussion

A total of 165 differences were generated by 80 participants across all four conditions: 35 in the separate interpretation condition, 37 in the joint interpretation condition, 32 in the Joint + Corr condition, and 61 in the control condition. Table 2 shows examples of the differences generated.

Rating the Differences

Scoring Procedure: A randomized list of the difference responses was rated by two laboratory assistants. They had no prior knowledge of this study and were blind to the conditions of the participants. Two different kinds of ratings were collected: (a) the causal relevance of the differences and (b) whether the differences were alignable. For the causal relevance ratings, each rater judged the quality of each difference in terms of its depth of causal relevance to heat transfer. The rating scale allowed respondents to assign a rating of 0 for (*not relevant*) or a rating from 1 (*shallow*) to 7 (*deep*). To help the raters formulate their judgments of causal relevance, we provided a set of key concepts (heat source, heat conductor, phase change) to be used in evaluating the responses.

For the alignability ratings, the criteria used to code each difference as alignable or nonalignable were adapted from the conservative technique used in Gentner and Markman (1994). A listed difference was counted as an alignable difference if (a) the participant mentioned contrasting properties of the two items (e.g., "A hotel is expensive; a motel is cheap") or (b) the participant used an explicit comparative construction (e.g., "Dogs are bigger than cats"). All other differences were considered nonalignable differences, including simple negation of one item's property as

TABLE 2
Experiment 1: Sample Responses Generated in the Difference Listing Task

Differences rated as "deep, interesting"
- In A pancakes are being solidified, but in B ice cube is liquifying, even though heat is applied and transferred in both situations. (Joint + Corr condition)
- The end product for the transfer of heat for A is cooked pancakes, while in B it is melted ice or water. (Joint + Corr condition)
- Heat is being transferred directly in A and indirectly in B. (Joint interpretation condition).
- Ice is liquifying, pancakes are solidifying. (across conditions)

Differences rated as "shallow" or "extraneous"
- You can eat the pancakes and drink the coffee. (Separated interpretation condition)
- One is being heated, the other is previously heated. (Separated interpretation condition)
- B is vertical and A is horizontal. (Control condition)
- Situation A is probably occurring in the southern states, situation B in the north. (Control condition)

applied to the other (e.g., "A police car has weapons in it, an ambulance does not"). As such, each difference was coded as either alignable or nonalignable.

Each rater independently scored each item and then the two raters discussed their judgment to resolve discrepancies. The initial independent causal relevance ratings showed 74% identical scores on the (0–7) scale (and 96% within a range of + or − 1). For a given participant's listed differences, the causal relevance ratings were combined to yield the mean score. The initial alignability coding showed 95% agreement and yielded the number of alignable and nonalignable differences for that participant.

The causal relevance rating results of differences.   The results of the causal relevance rating are shown in Figure 2. All analyses were conducted using the conservative technique of an omnibus test supported by post hoc comparisons. As predicted, there was a difference across conditions. A one-way, between-subject analysis of variance (ANOVA; 4 levels) using the mean causal rating of each participants differences as the dependent measure showed an overall significant difference, $F(3, 76) = 2.75, p < .05$. A Dunnett test showed a significant difference between the Joint + Corr and control conditions ($p < .05$). No other pairwise comparisons reached significance levels.

The rating results support the claim that mutual alignment can be effectively induced by the combined tasks of joint interpretation and correspondence. Although the presented scenarios differed obviously in terms of the objects involved and the thematic context, comparison guided participants toward listing differences causally related to the common theme of heat transfer. This result suggests the efficacy
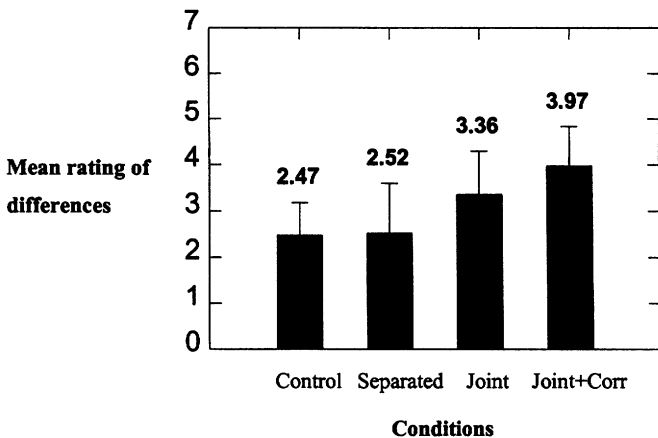
FIGURE 2   Experiment 1: Mean rating of differences in terms of the causal relevance to heat transfer.

of mutual alignment as a bootstrapping process that can facilitate better understanding of the presented examples by emphasizing the encoding of the common relational structure. The findings also suggest that mutual alignment is sensitive to the nature of the comparison context in the orienting task. In the separate condition, providing participants with the opportunity to compare through sequential presentation and description of scenarios appears to be nearly the same as the control condition of no orienting task at all. The differences listed by participants were rated as more causally relevant to heat flow when they performed a comparison task of joint interpretation, but the effect is reliable only in the Joint + Corr condition.

The alignability coding results of differences.   The results of our secondary scoring approach based on the conservative technique for coding differences for alignability are shown in Table 3A. Successful alignment is reflected in the proportion of ADs to the total differences (AD / [NAD + AD]). Results are consistent with the claim that an effective comparison led to a higher proportion of ADs, however the finding failed to reach levels of significance, $F(3, 76) = 1.33, p > .10$. This conservative scoring procedure is likely to underestimate the phenomenon, so the lack of a reliable difference across conditions may speak to the stringency of this methodological approach rather than the manipulation itself.

Rating the Descriptions

Scoring procedure.   A randomized list of all participants' descriptions of the coffee cup scenario was rated by the same two laboratory assistants to assess understanding of the scenarios and test for re-representation effects (see Table 4 for

TABLE 3
Experiment 1: Overview of Mean Scores Results On All Dependent Measures

Conditions	Joint + Corr (n = 20)	Joint (n = 20)	Separated (n = 20)	Control (n = 20)
A. Differences				
Causal relevance rating of differences (0–7)	3.97* (1.85)	3.36 (2.00)	2.52 (2.31)	2.47 (1.51)
Proportion of ADs	0.72 (0.40)	0.59 (0.44)	0.53 (0.43)	0.48 (0.27)
B. Descriptions				
Causal relevance rating of descriptions (0–7)	6.3 (0.57)	6.0 (1.34)	5.70 (1.56)	5.25 (1.77)
C. Similarity				
Rating of similarity (1–7)	4.95* (1.15)	4.05 (1.43)	3.85 (1.50)	3.50 (1.15)

Note.   ADs = alignable differences.
*Significantly different from control.

"Deep, clear" description
- Heat is transferred from the ice cube through the metal bar, causing the ice to melt. (across conditions)
- The hot coffee heats the metal bar, which in turn transfers heat to the ice cube causing it to melt. (Joint + Corr condition)

"Shallow" description
- A metal rod with an ice cube at the end is in coffee. I do not know why. (Control condition)
- A man ordered a cup of coffee. It was too hot. We assume it was a man in his late 30s, and he took out the 12" metal bar he kept in his pocket and speared an ice cube, so that the ice would not melt into his coffee. (Separated interpretation condition)

examples of the descriptions generated). Each rater was asked to evaluate the quality of each description in terms of depth of causal relevance to heat transfer on a scale from 0 (*low*) to 7 (*high*). Each rater independently scored a given description, and then the raters discussed the results. The initial scoring showed 92% agreement and all discrepancies were resolved by discussion.

The causal relevance rating results of descriptions.   The mean ratings of the descriptions in terms of depth of causal relevance to heat flow are shown in Table 3B. Although participants received no instruction (explicit or implicit) to describe the coffee cup scenario in terms of the heat transfer schema, many included it in their description under comparison conditions. The Joint + Corr group showed the best performance ($M = 6.3$) of all the conditions. Although we did not find significant differences across conditions, $F(3, 76) = 2.09, p \approx .10$, there is a clear trend in the predicted direction. These results are consistent with our specific predictions for ordering among the orienting tasks and provide another source of convergent evidence for the role of mutual alignment in fostering understanding of the common structure of the two scenarios.

The failure to reach significance levels in this measure may reflect a lack of sensitivity of the description measure or the scoring procedure. In examining the description results, we discovered that participants who performed a comparison (joint interpretation and Joint + Corr conditions) were more likely to use terms referring to the causal schema of heat transfer (heat, transfer, and cause) than those who did not (separate interpretation and control conditions). After counting the number of participants who used these terms, we found significant differences between groups (see Table 5). For heat, $\chi(3) = 9.43, p < .05$. For cause, $\chi(3) = 10.41$, $p < .05$. However, for perceptual and descriptive words specific to the coffee cup or ice cube objects (melt, hot, or cold) no such difference was found. This difference in use of the schema-relevant terms suggests a richer and more uniform repre-

TABLE 5
Experiment 1: Number of Participants Using Key Terms in Scenario Descriptions

	Conditions	*Joint + Corr* *(n = 20)*	*Joint* *(n = 20)*	*Separated* *(n = 20)*	*Control* *(n = 20)*
Terms referring	Heat	19	18	16	12
to the	Cause (because of, due to)	16	18	10	11
common schema of heat transfer	Transfer (conduct, transmit)	12	14	13	9
Perceptual and	Melt	19	16	15	16
descriptive words	Hot (cold)	10	15	15	16

sentation of the scenarios. To further evaluate the induction or instantiation of a heat-transfer schema we checked for use of multiple schema-relevant terms in scenario descriptions. We found that 48% of participants in the joint interpretation conditions used all three key terms (heat, transfer, and cause) in their descriptions as opposed to 23% of participants in the noncomparing conditions. The tendency for the comparison groups to invoke elements of a causal schema supports our view of comparison promoting highlighting and abstraction of common structure along with increased representational uniformity.

The similarity judgment results.   The mean ratings of similarity given by the participants are shown in Figure 3. The pattern is the same as those found in the difference task and the description task. ANOVA shows overall significant differences, $F(3, 76) = 4.42, p < .05$. Furthermore, Dunnett tests show significant difference between Joint + Corr and control condition $(p < .05)$, and Joint + Corr and separate interpretation condition $(p < .05)$. No other differences were significant. Participants in the Joint + Corr condition judged the scenarios to be more similar than did noncomparing participants with or without an orienting task.

The ratings of descriptions and the similarity judgments are consistent with the results from the difference-listing task. The Joint + Corr orienting task appears to promote better alignment and comprehension as measured in the difference listing task, word use in the description task, and the similarity rating task. In contrast, joint interpretation alone nor separate interpretation reliably lead to improved performance relative to baseline.

To further look into the relation among the dependent measures, we obtained Pearson correlations and probabilities as shown in Table 6. We took the proportion of AD as a measure of alignment across any kind of commonalities and the causal relevance rating of differences as an indirect measure of alignment along the common causal structure of primary interest. As predicted, these measures were strongly
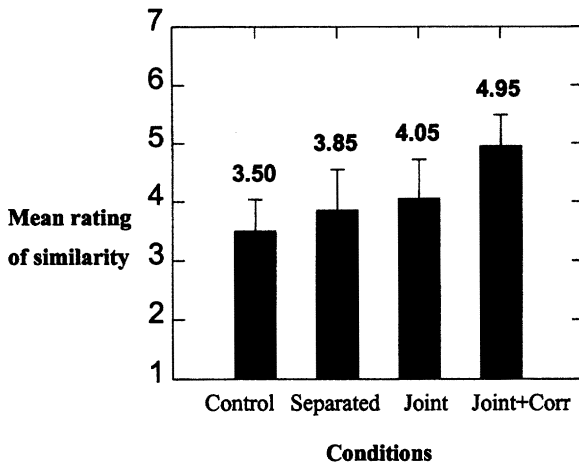
FIGURE 3   Experiment 1: Mean rating of similarity.

TABLE 6
Experiment 1: Results of Pearson Correlation and Probabilities Matrix

r	Proportion of AD	Causal Rating of Differences	Causal Rating of Descriptions	Rating of Similarity
Causal rating of differences	0.68**			
Causal Rating of descriptions	0.33**	0.30**		
Rating of similarity	0.28**	0.28**	0.12	
Science background	−0.05**	0.03**	0.07	0.20

Note.   Number of paired observations was 80 per cell. AD = alignable difference.
*p < .05. **p < .01.

correlated. We also found positive correlations between the causal relevance ratings of the scenario descriptions and the two difference scores with the proportion of AD and the ratings of causal relevance of the differences. These convergent results lend additional support to the claim that mutual alignment promoted insight into deeper common structure. Additionally, the positive correlation between the similarity ratings and the two difference measures is consistent with Gentner and Markman's (1994) finding that the number of alignable differences listed by participants tends to increase with the degree of similarity of the items.

Finally, as shown in Table 6, participants' responses in the measurement tasks were not correlated to their self-rated science backgrounds. Therefore, different interpretations by the participants are not simply a reflection of their previous knowledge, but instead are attributable to the experimental manipulation.

EXPERIMENT 2

In Experiment 1, we found that participants in the Joint + Corr condition performed better than the other groups across a series of dependent measures. However, because two types of orienting tasks were involved in Joint + Corr condition, it is not clear whether it is one of the tasks or the combination that accounts for the finding of effective mutual alignment. In particular, because the joint interpretation (alone) group did not perform as well, it might be argued that the correspondence task itself was the driving force behind our pattern of data. We conducted the following experiment to test whether the correspondence task would promote alignment without a prior joint interpretation task. In the experiment, we divided participants into two groups. One group only performed a correspondence task, whereas the other group performed joint interpretation plus correspondence task as in Experiment 1. The accuracy of the responses generated in the correspondence task serves as a measure of alignment.

For scenarios that are structurally similar, the correspondence task can be challenging. Lacking a good alignment, participants may attend to the superficially similar elements and miss the relational matches. Only good alignment enables the participants to assign correct correspondences between elements of the two scenarios. Because joint interpretation provides a context that invites participants to focus on the common structure and relations, we expect it to promote alignment and facilitate the correspondence task. Therefore, the joint interpretation plus correspondence group is expected to give more correct responses than the correspondence-only group.

METHOD

Participants

Twenty undergraduate students at Northwestern University fulfilling a course requirement participated in the study. Ten participants were randomly assigned to each condition.

Materials

The scenarios and task sheets from Experiment 1 were used.

Design and Procedure

There were two conditions run in a between-subject design. The experimenter gave each participant the task sheets and asked them to fill out the task sheets in order.

The joint interpretation plus correspondence group started with the joint interpretation task followed by the correspondence task. The correspondence-only group were given only the correspondence task.

Results and Discussion

Scoring procedure.   Participants' responses in the correspondence task were scored by the experimenter. The correct answers are listed in Table 7. We adopted these answers according to the following alignment: Heat is transferred from the gas stove (coffee cup) through the griddle (metal bar), causing the batter (ice cube) to become pancakes (drops of water).

The difference between the mean number of correct answers (see Figure 4) given by the participants in the joint interpretation plus correspondence condition and the correspondence-only condition was reliable, $t(9) = 2.684$, $p < .05$. In Experiment 1, there were also 20 participants in the joint interpretation plus correspondence condition. To verify the observed pattern, we scored their correspondence responses and obtained comparable results. The mean number of correct responses was 2.95, which is nearly identical to the mean result in the joint interpretation plus correspondence condition in this study.

This result provides evidence that the correspondence task itself may not be sufficient to produce a quality conceptual alignment. Experiment 1 suggests that the correspondence task and joint interpretation task can be used in conjunction to help learners focus on the common structure and relational matches. We conclude that both the joint interpretation and the correspondence tasks play a causal role and that the two in combination form an effective basis for learning by mutual alignment.

GENERAL DISCUSSION

Our principle goal has been to demonstrate mutual alignment as a basis for learning. When participants were led to carry out joint interpretation and correspondence, they showed better causal understanding. Compared to the control

TABLE 7
Experiment 2: Correct Correspondences Between Situation A and Provided Elements of Situation B

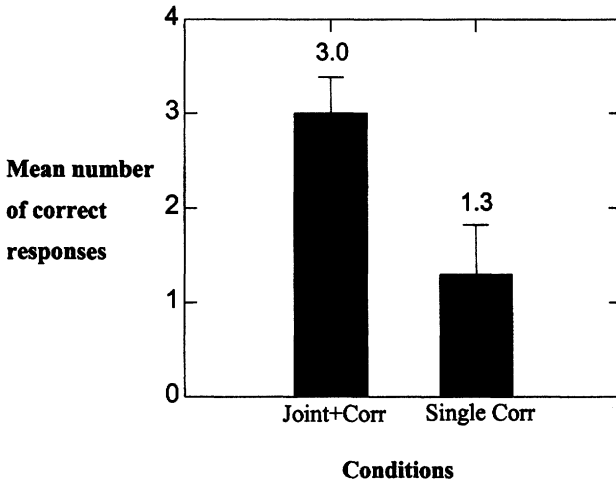| Situation A	Situation B
1. Batter (or pancake)	1. Ice cube
2. Griddle (pan)	2. Metal bar
3. Gas stove (or flame)	3. Coffee
4. Pancake	4. Drops of water

FIGURE 4   Experiment 2: Mean number of correct responses in the correspondence task.

condition, this group generated differences more relevant to the causal structure of heat transfer and judged the scenarios to be more similar. The deeper understanding of the stimulus materials by participants in the joint interpretation and correspondence condition is evidenced by reliably better performance in

1. Noticing insightful differences between the scenarios.
2. Using key terms relevant to the heat-transfer schema to describe the scenarios.
3. Appreciating the similarity between the two scenarios.

The evidence provides support for the claim that intensive comparison of two analogous situations leads to mutual alignment and improved encoding of the examples in terms of the common higher order relation of heat flow.

In Experiment 1 we found evidence for two components of successful mutual alignment: joint interpretation and correspondence. Inviting participants to consider two scenarios together promoted a focus on the common structure and relations. The second component of successful mutual alignment was the correspondence task. In the process of assigning one-to-one correspondences between scenarios, participants had to consider the entire system and make explicit the parallels between elements. We did not see evidence in Experiment 2 that correspondence by itself could lead to enhanced understanding. Rather, as expected, engaging learners in more active comparison processes—involving both correspondence and joint interpretation—proved more efficacious.

Re-representation during comparison has been posited as a mechanism that can lead to common higher order abstractions. We speculate that re-representation

processes may be responsible in part for the higher similarity ratings by the comparison groups. This result is consistent with the findings of Kurtz (1996) and Livingston, Andrews, and Harnard (1998) that category learners judged items from the same category as more similar than did control groups who did not acquire the category structure. Linking examples together by learning their category structure, by successful alignment, or by a combination of these factors may alter the way in which the individual examples are understood. Another line of support for re-representation is the tendency toward convergent word use in descriptions under the joint interpretation conditions. This suggests the possibility that the juxtaposition of scenarios increased the uniformity of the representations.

Limitations of This Research

One concern in interpreting these results is whether differences in amount of processing time during the orienting task could be responsible for the pattern of performance observed across conditions. The entire procedure was self-paced, and we did not time the task. Thus, we cannot rule out the possibility that the tasks that best promoted learning were those on which learners spent the most time. For example, because there was no orienting task in the control condition, poor performance here could be attributable to spending less overall time processing the stimulus materials. We note, however, that while the intensity of the comparison context increased progressively across the orienting tasks, this might not be the case for time-on-task. The separate interpretation group had to write two descriptions, whereas the joint interpretation group only had to write one. Yet, the data in these conditions follow the pattern predicted by the nature of the task, not the amount of time required. Along these lines, the correspondence listing task (which in combination with joint interpretation resulted in successful mutual alignment) is particularly demanding in terms of the depth or specificity of comparison required, but it is typically performed fairly quickly. Again, the contribution of the correspondence listing is more easily attributable to the type of processing than to the time spent on task. Overall, we believe that understanding in this task derives from the right kind of processing, rather than merely amassing processing time. However, it would be useful in future research to address this issue by equating time-on-task across conditions.

　　Although the present results provide promising initial evidence, many questions remain. The generality of our findings is limited, because only one set of stimulus materials was tested. In addition, the scenarios we used were self-contained and depicted relatively straightforward conceptual content. In real-world experience, cases tend to be more complex. In the classroom, cases are often presented in rich contexts—as part of a lecture, discussion, or exercise. The effects of mutual alignment within our simple experimental context were encouraging, but not dramatic. However, we note that, although scaling up is typically a

challenge, it is possible in this case that more intricate materials would reveal larger effects, because mutual alignment should be most effective in learning domains rich in relational structure. For example, Loewenstein et al. (1999) found robust effects of analogical encoding using complex, verbal scenarios. Another basis for cautious optimism about the potential generality of our findings is that structural alignment theory has proven applicable to a broad set of domains: verbal and pictorial, simple and complex. Further research will be necessary to see whether and how this learning technique will extend to a broader range of materials and contexts.

Implications for Education

The patterns revealed in this study have implications for education and training. In the practice of explaining concepts by analogy, teachers typically give vivid examples as the information base and encourage the students to transfer these examples to solve the problems. However, even when these examples are potentially useful, students may not be able to make use of the analogy between a new problem and the example if they are hindered by a superficial distinction (Lancaster & Kolodner, 1987; Novick, 1988; Ross, 1989). According to Catrambone, Jones, Jonides, and Seifert (1995), when solving a problem that reminds them of a familiar example, people often reason by analogy to the instance. However, as noted earlier, instances are often retrieved on the basis of superficial similarity (Gentner, Rattermann, & Forbus, 1993; Holyoak & Koh, 1987; Ross, 1989). Thus, even though learners may be capable of sophisticated analogical reasoning, they may be limited to finding superficial commonalities because the retrieved example contains nothing better.

Our results suggest that learning can be promoted by inviting comparison. Inducing learners to evaluate analogous situations jointly can allow for the discovery of common structure and help people generate a retrieval cue for later use. This is consistent with the suggestion by case-based reasoning researchers (Kolodner, Narayanan, & Hmelo, 1996; Narayanan & Kolodner, 1995; Schank, Fano, Bell, & Jona, 1993) that libraries of relevant cases be made available to students as they are solving problems, because such cases can suggest issues to focus on and solutions to problems (Kolodner, 1997). In addition, mutual alignment emphasizes bidirectional processing of partially understood examples, rather than the more traditional approach of analogical transfer from a well-understood base to a poorly understood target. Finally, our findings underscore the importance of going beyond mere availability or shallow processing of multiple examples. It is important to encourage learners to articulate mappings; for example, to lay out the full set of correspondences. Clearly articulating the correspondences requires learners identify the objects roles in the global system and label the objects at an abstract level.

Mutual alignment as formulated in this research can be translated into the classroom context. In the context of lecture or reading, case-based learning can be extended to the presentation of multiple examples along with an active comparison context. Mutual alignment lends itself naturally to active and participatory learning approaches as the comparison process requires an active stance rather than passively receiving information. Beyond carrying out comparison exercises during instruction or at home, another approach is to divide students into two groups, each working on different, but conceptually related, problems or domains during the activity. The groups can be brought together to interpret the problems jointly and work out the correspondences. Such implementations of mutual alignment may lead to immediate or secondary consequences in classroom learning. Schwartz and Bransford (1998) found that the benefits of analyzing contrasting cases were revealed in subsequent instruction—the technique was most effective in enabling learning from traditional modes such as reading texts or hearing lectures. Thus, mutual alignment may hold promise as a scaffolding technique within a broader pedagogical approach.

Additional implications of these findings bear consideration. Mutual alignment may be helpful in addressing the inert knowledge problem: Learners often have appropriate knowledge, but fail to use it when it is relevant (Bransford, Franks, Vye, & Sherwood, 1989). Although the knowledge is applicable in many situations, learners tend to be reminded only in a restricted context and in a narrow domain. Mutual alignment can bypass the difficulty of activating prior analogs by encouraging learners to induce schemas that may help during subsequent problem solving and improve the quality of instruction and learning (Faries & Reiser, 1988; Ferguson, Bareiss, Birnbaum, & Osgood, 1992; Seifert, 1989).

Along related lines, learners often make a distinction between everyday life, viewed as a "real-world" situation, and scientific knowledge, viewed as abstract laboratory knowledge (Clement, 1993; diSessa, 1983). For example, in our study, the unusual situation of a coffee cup with an ice cube in it was described by participants as a scientific experiment, whereas the familiar situation of cooking pancakes was understood as a real-world example. From this viewpoint, people may be unaware that seemingly distant situations have matching underlying structures (Catrambone et al., 1995; Kaiser, Jonides, & Alexander, 1986). It can be hard to help students comprehend the general principles underlying both a science experiment and everyday life experiences given the variation in context and surface details. For example, in the Kaiser et al. study, participants predicted a straight path for water leaving a tube but a curved path for a ball, although both cases are equivalent in physics principles. By encouraging learners to compare everyday situations and scientific situations, mutual alignment may promote deeper understanding and appreciation of the operation of abstract principles in the real world.

Mutual alignment may be helpful in early stages of domain learning. Although novice learners may be limited to highly similar cases with large amounts of overlapping information, there is evidence that even a relatively easy alignment may make possible more abstract analogies based on the same relational structure. Using a perceptual similarity task, Kotovsky and Gentner (1996) showed that 4-year-old children given concentrated experience comparing highly similar, same-dimension figures (that could easily be aligned) were then able to process otherwise non-obvious cross-dimensional matches. Thus experiencing close alignments allowed children to subsequently recognize abstract relational similarity (e.g., a common pattern across the size dimension and the color dimension), whereas children without this comparison experience did not show such recognition. This progressive alignment process may help early learners move from concrete representations to more abstract representations (see Figure 5). A related technique is Clement's (1993) use of bridging analogies to move the learner from relative close similarities to a more abstract and less transparent analogy.

Despite these various possibilities, there are limitations on the application of the mutual alignment approach. Mutual alignment can enhance understanding, but there must be at least rudimentary prior knowledge to make two-way information transfer possible. Learning by mutual alignment can succeed without analogical retrieval, but it is not bootstrapping from ground zero. Mutual alignment also depends on enticing the learner into an active process by providing a learning context that fosters deep comparison. Further, if the goal is for a learner to be able to interpret single target examples using analogical reasoning, mutual alignment serves only as a learning tool toward this ability, not a basis for it in its own right.

Further Research

This discussion of implications and limitations of our findings raises the question of directions for further research. A top priority is to test the effects of mutual alignment beyond improving the understanding and encoding of training materials. To test participants' ability to apply what they have learned after mutual alignment, they should be tested on their ability to transfer their understanding to novel cases. Another approach would involve giving participants a series of questions designed to test the quality and limits of their understanding after mutual alignment. Are difficult or unusual examples more effectively understood after mutual alignment of standard examples?

Another direction for future work is to look more closely at the detailed operation of the comparison process. These data do not fully pinpoint the role of specific mechanisms acting as component processes within alignment: highlighting, infer-
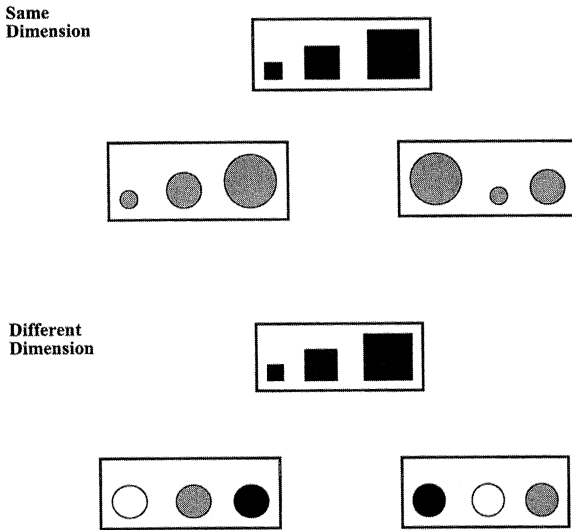
Same
Dimension

Different
Dimension

FIGURE 5   Example of materials used in Kotovsky and Gentner (1996). From "Comparison and Categorization in the Development of Relational Similarity," by L. Kotovsky and D. Gentner, 1996, *Child Development, 67*, pp. 2797–2822. Copyright 1996 by Society for Research in Child Development, University of Michigan. Re-used with permission.

ence, re-representation, and abstraction. To evaluate re-representation, it would have been useful to compare initial representations with later ones. However, in this design, obtaining initial representations of the coffee cup scenario would have interfered with the orienting tasks. Further research is underway to address the role of re-representation in the comparison process.

A host of additional questions deserve consideration. How does mutual alignment compare to other techniques for promoting learning? Can tasks other than joint interpretation plus correspondence listing be used to implement mutual alignment? How effective is mutual alignment with different types of cases including verbal materials or more complex materials? How does the effectiveness of mutual alignment vary with the similarity of the two analogs (e.g., literally similar or highly remote analogs)? What level of prior knowledge is optimal for mutual alignment? Is the technique suitable for novices or for advanced learners? Is it possible to speed up the path to expertise via mutual alignment?

SUMMARY

Analogy is a powerful tool for learning and reasoning. However, a good analogy is often difficult for learners to construct or retrieve. We demonstrate the effi-

cacy of mutual alignment as a technique for fostering learning by intensive comparison of two partially understood examples. Our results provide evidence about the conditions under which comparison fosters analogical insight by revealing or highlighting the common structure between examples. We consider this to be a kind of bootstrapping process as improved understanding of the examples is fostered without the need to find or draw on well-understood cases or sophisticated knowledge structures from memory. Mutual alignment by joint interpretation plus correspondence listing promotes richer encoding and understanding in the laboratory setting—the technique may prove beneficial in facing broad, real-world challenges in teaching and learning.

ACKNOWLEDGMENTS

REFERENCES

Bransford, J., Franks, J., Vye, N., & Sherwood, R. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470–497). New York: Cambridge University Press.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 1147–1156.

Catrambone, R., Jones, C. M., Jonides, J., & Seifert, C. (1995). Reasoning about curvilinear motion: Using principles or analogy. *Memory & Cognition, 23,* 368–373.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science, 15,* 89–132.

Clement, C. A., Mawby, B., & Giles, D. E. (1994). The effects of manifest relational similarity on analog retrieval. *Journal of Memory and Language, 33,* 396–420.

Clement, J. (1993). Using analogies to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30,* 1241–1257.

Cummins, D. D. (1992). Role of analogical reasoning in the induction of problem categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18,* 1103–1124.

diSessa, A. A. (1983). Phenomenology and the evolution of intuition. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 15–33). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17,* 391–434.

Dunbar, K. (1995). How scientists really reasoning: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365–395). Cambridge, MA: MIT Press.

Falkenhainer, B. (1990). Analogical interpretation in context. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society* (pp. 69–76). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial intelligence, 41,* 1–63.

Faries, J. M., & Reiser, B. J. (1988). Access and use of previous solutions in a problem solving situation. In *Proceedings of the 10th Annual Conference of the Cognitive Science Society* (pp. 433–439). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ferguson, R.W. (1994). MAGI: A model of analogy-based encoding using symmetry and regularity. In *Proceedings of the Sixteenth Annual Conference on the Cognitive Science Society* (pp. 283–288). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ferguson, R. W., & Forbus, K. D. (1998). Telling juxtapositions: Using repetition and alignable difference in diagram understanding. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research* (pp. 109–117). Sofia, Bulgaria: New Burlingame University.

Ferguson, W., Bareiss, R., Birnbaum, L., & Osgood, R. (1992). ASK systems: An approach to the realization of story-based teachers. *Journal of the Learning Sciences, 2,* 95–134.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19,* 141–205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155–170.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical learning* (pp. 199–241). London: Cambridge University Press.

Gentner, D., Brem, S., Ferguson, R. W., Markman, A. B., Levidow, B. B., Wolff, P., & Forbus, K. D. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *The Journal of the Learning Sciences, 6,* 3–40.

Gentner, D., & Gunn, V. (in press). Structural alignment facilitates the noticing of differences. *Memory and Cognition.*

Gentner, D., & Holyoak, K. J. (1997). Reasoning and learning by analogy: Introduction. *American Psychologist, 52,* 32–34.

Gentner, D., & Kurtz, K. J. (2001). *Mechanisms of mapping: Evidence from online analogy evaluation.* Manuscript in preparation.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science, 5,* 152–158.

Gentner, D., & Markman, A. B. (1995). Similarity is like analogy. In C. Cacciari (Ed.), *Similarity* (pp.111–147). Brussels, Belgium: BREPOLS.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52,* 45–56.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65,* 263–297.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14,* 487–513.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: 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 retrieval 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.

Gentner, D., & Wolff, P. (2000). Metaphor and knowledge change. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual change in humans and machines* (pp. 295–342). Cambridge, MA: MIT Press.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

Goldstone, R. L. (1994a). The role of similarity in categorization: Providing a groundwork. *Cognition, 52*(2), 125–157.

Goldstone, R. L. (1994b). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 3–28.

Goldstone, R. L., & Medin, D. L. (1994). The time course of comparison. *Journal of experimental psychology: Learning, Memory and Cognition, 20,* 29–50.

Goldstone, R. L., & Medin, D. L., & Gentner, D. (1991). Relations attributes and the non-independence of features in similarity judgment. *Cognitive Psychology, 23,* 222–264.

Hesse, M. B. (1966). *Models and analogies in science.* Notre Dame, IN: University of Notre Dame Press.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 19). New York: Academic.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15,* 332–340.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295–355.

Holyoak, K. J., & Thagard, P. (1995). *Mental leap: Analogy in creative thought.* Cambridge, MA: MIT Press.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52,* 35–44.

Kaiser, M. K., Jonides, J., & Alexander, J. (1986). Intuitive reasoning about abstract and familiar physics problems. *Memory & Cognition, 14,* 308–312.

Keane, M. T. (1988). Analogical mechanisms. *Artificial Intelligence Review, 2,* 229–250.

Kolodner, J. (1993). *Case-based reasoning.* San Mateo, CA: Morgan Kaufmann.

Kolodner, J. L. (1997). Educational implications of analogy: A view from case-based reasoning. *American Psychologist, 52,* 57–66.

Kolodner, J. L., Narayanan, H., & Hmelo, C. (1996). Problem-based learning meets case-based reasoning. In E. Domeshek & D. Edelman (Eds.), *Proceedings of the 1996 International Conference of the Learning Sciences* (pp. 188–195). VA: American Association for Computers in Education.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67,* 2797–2822.

Kurtz, K. J. (1996). Category-based similarity. In *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (p. 790). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kurtz, K. J., Gentner, D., & Gunn, V. (1999). Reasoning. In B. Martin Bly & D. E. Rumelhart (Eds.), *Handbook of perception and cognition: Cognitive science* (pp. 144–200). San Diego, CA: Academic.

Lancaster, J. S., & Kolodner, J. L. (1987). Problem solving in a natural task as a function of experience. In *Proceedings of the Ninth Annual Conference of the Cognitive Science society* (pp. 715–726). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Livingston, K. R., Andrews, J. K., & Harnard, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology, 24,* 732–753.

Loewenstein, J., Gentner, D., & Thompson, L. (in press). *Analogical encoding and relational language in long-term transfer of negotiation strategies.*

Loewenstein, J., Thompson, L., Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6,* 586–597.

Markman, A. B., & Gentner, D. (1993a). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and 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. *Psychological Science, 8,* 363–367.

Markman, A. B., & Medin, D. L. (1995). Similarity and alignment in choice. *Organizational Behavior and Human Decision Processes, 63,* 117–130.

Mayer, R. E. (1992). *Thinking, problem solving, cognition.* San Francisco: Freeman.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278.

Narayanan, H., & Kolodner, J. L. (1995). Case libraries in support of design education: The DESIGNMuse experience. In *Proceedings FIE 5: Frontiers in education* (pp. 2b2.1.2). Atlanta, GA: IEEE Press.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere & H. Feigl (Eds.), *Cognitive models of science: Minnesota studies in the philosophy of science* (pp. 3–44). Minneapolis, MN: University of Minnesota Press.

Novick, L. R. (1988). Analogical transfer, problem similarity and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 510–520.

Novick, L. R. (1990). Representational transfer in problem solving. *Psychology Science, 1,* 128–132.

Novick, L. R., & Hmelo, C. E. (1994). Transferring symbolic representations across nonisomorphic problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20,* 1296–1321.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 398–415.

Perfetto, G. A., Bransford, J. D., & Franks, J. J. (1983). Constraints on access in a problem solving context. *Memory & Cognition, 11,* 24–31.

Pirolli, P. L., & Anderson, J. R. (1985), The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology, 39,* 240–272.

Polya, G. (1945). *How to solve it.* Princeton, NJ: Princeton University Press.

Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 124–139.

Reeves, L. M., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin, 3,* 381–400.

Riesbeck, C. K., & Schank, R. C. (1989). *Inside case-based reasoning.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16,* 371–416.

Ross, B. H. (1986). Remindings in learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity, analogy and thought* (pp. 438–469) New York: Cambridge University Press.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 629–639.

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, and Cognition, 15,* 456–468.

Rumelhart, D. E. (1989). Toward a micro-structural account of human reasoning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical learning* (pp. 298–312). London: Cambridge University Press.

Schank, R. C., Fano, A., Bell, B., & Jona, M. (1993). The design of goal-based scenarios. *Journal of the Learning Sciences, 3,* 305–346.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition & Instruction, 16,* 475–522.

Seifert, C. M. (1989). Analogy and case-based reasoning. In K. J. Hammond (Ed.), *Proceedings: Second workshop on case-based reasoning* (DARPA; pp. 125–129). San Mateo, CA: Morgan Kaufmann.

Sternberg, R. J. (1977). *Intelligence, information processing and analogical reasoning: The componential analysis of human abilities.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Thagard, P. (1992a). *Conceptual revolutions.* Princeton, NJ: Princeton University Press.

Thagard, P. (1992b). Analogy, explanation, and education. *Journal of Research in Science Teaching, 29,* 537–544.

Thompson, L., Gentner, D., & Loewenstein, J. (2000). Avoiding missed opportunities in managerial life: Analogical training more powerful than individual case training. *Organizational Behavior & Human Decision Process, 82,* 60–75.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

VanderStoep, S. W., & Seifert, C. M. (1993). Learning "how" versus learning "when:" Improving transfer of problem-solving principles. *Journal of the Learning Sciences, 3,* 93–111.

Weisberg, R., DiCamillo, M., & Phillips, D. (1978). Transferring old associations to new situations: A nonautomatic process. *Journal of Verbal Learning Behavior, 17,* 219–228.

You have printed the following article:

Learning by Analogical Bootstrapping
Kenneth J. Kurtz; Chun-Hui Miao; Dedre Gentner
The Journal of the Learning Sciences, Vol. 10, No. 4. (2001), pp. 417-446.
Stable URL:
http://links.jstor.org/sici?sici=1050-8406%282001%2910%3A4%3C417%3ALBAB%3E2.0.CO%3B2-E

*This article references the following linked citations. If you are trying to access articles from an
off-campus location, you may be required to first logon via your library web site to access JSTOR. Please
visit your library's website or contact a librarian to learn about options for remote access to JSTOR.*

References

ASK Systems: An Approach to the Realization of Story-Based Teachers
William Ferguson; Ray Bareiss; Lawrence Birnbaum; Richard Osgood
The Journal of the Learning Sciences, Vol. 2, No. 1. (1992), pp. 95-134.
Stable URL:
http://links.jstor.org/sici?sici=1050-8406%281992%292%3A1%3C95%3AASAATT%3E2.0.CO%3B2-L

Analogical Reasoning and Conceptual Change: A Case Study of Johannes Kepler
Dedre Gentner; Sarah Brem; Ronald W. Ferguson; Arthur B. Markman; Björn B. Levidow; Phillip
Wolff; Kenneth D. Forbus
The Journal of the Learning Sciences, Vol. 6, No. 1, Conceptual Change. (1997), pp. 3-40.
Stable URL:
http://links.jstor.org/sici?sici=1050-8406%281997%296%3A1%3C3%3AARACCA%3E2.0.CO%3B2-%23

Comparison and Categorization in the Development of Relational Similarity
Laura Kotovsky; Dedre Gentner
Child Development, Vol. 67, No. 6. (Dec., 1996), pp. 2797-2822.
Stable URL:
http://links.jstor.org/sici?sici=0009-3920%28199612%2967%3A6%3C2797%3ACACITD%3E2.0.CO%3B2-1

The Design of Goal-Based Scenarios
Roger C. Schank; Andrew Fano; Benjamin Bell; Menachem Jona
The Journal of the Learning Sciences, Vol. 3, No. 4, Goal-Based Scenarios. (1993 - 1994), pp. 305-345.
Stable URL:
http://links.jstor.org/sici?sici=1050-8406%281993%2F1994%293%3A4%3C305%3ATDOGS%3E2.0.CO%3B2-6

A Time for Telling
Daniel L. Schwartz; John D. Bransford
Cognition and Instruction, Vol. 16, No. 4. (1998), pp. 475-522.
Stable URL:
http://links.jstor.org/sici?sici=0737-0008%281998%2916%3A4%3C475%3AATFT%3E2.0.CO%3B2-6

Learning "How" versus Learning "When": Improving Transfer of Problem-Solving Principles
Scott W. VanderStoep; Colleen M. Seifert
The Journal of the Learning Sciences, Vol. 3, No. 1. (1993 - 1994), pp. 93-111.
Stable URL:
http://links.jstor.org/sici?sici=1050-8406%281993%2F1994%293%3A1%3C93%3AL%22VL%22I%3E2.0.CO%3B2-7