UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
The Influence of Structural Support on the Evaluation of Analogical Inferences

Permalink
https://escholarship.org/uc/item/15f5x1pp

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
Laux, Jeffrey P.
Markman, Arthur B.

Publication Date
2006

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

The Influence of Structural Support on the Evaluation of Analogical Inferences

Jeffrey P. Laux (laux@mail.utexas.edu)
Department of Psychology, University of Texas, 1University Station A8000
Austin, TX 78712-0187 USA

Arthur B. Markman (markman@psy.utexas.edu)
Department of Psychology, University of Texas, 1University Station A8000
Austin, TX 78712-0187 USA

Abstract

Analogical inferences occur when knowledge about one domain is extended by virtue of its similarity to a second. How such inferences are evaluated is not yet known, but two factors (*support* and *extrapolation*) were suggested by Forbus, Gentner, Everett and Wu in 1997. We report an initial test of the role of support in inference evaluation. Subjects were asked to provide a confidence rating for an inference that had more or less structural support. The results yielded evidence that support influences the evaluation of inferences. This paper is discussed within the framework of the Structure-Mapping Theory of analogy.

Introduction

Analogy is a central cognitive process involved in representation, learning, categorization, and reasoning. An important use of analogies is to extend knowledge of one domain based on the analogy to a second domain. While some research has explored factors that govern how these inferences are made, little research has examined how they are evaluated. An important question that needs to be addressed is how people make judgments about the strength of an inference that is based on an analogy. Specifically, what factors make people think that an inference is more or less likely to be true? These questions will be addressed within the framework of Gentner's (1983) Structure-Mapping theory of analogy. There are other theories of analogy in cognitive psychology, but the general description used in the Structure-Mapping theory is compatible with the consensus of many researchers in the field (see, for example, Gentner, Holyoak, & Kokinov, 2001; Hummel, & Holyoak, 1997; Keane, Ledgeway, & Duff, 1994).

Review of Structure-Mapping Theory

Structure-Mapping Theory posits that concepts are represented "as propositional networks of nodes and predicates" (Gentner, 1983, p. 157). A predicate is a statement that is asserted of a subject or subjects; arguments are the subjects of which predicates are asserted. These predicates are partitioned into attributes (defined as taking single arguments) and relations (which take two or more arguments). For example, in the statement 'the boot is brown', *brown* is an attribute, and can be written using the one argument predicate Brown(boot). In the form 'Brown(boot)', 'Brown' is a predicate and 'boot' is its argument. The boot itself, outside of its role in the predicate, is an object. In the statement 'the boot is larger than the shoe', *larger-than* is a relation, which would be written with the two argument predicate Larger_than(boot,shoe). Relations can connect other relations as well as objects. The most common example of this is the relation *cause*, as in 'Bob is taller than Sam, causing Bob to jump higher than Sam', written Cause(Taller_than(Bob,Sam), Can_jump_higher_than(Bob, Sam)). Such a relation is known as a higher order relation (in the preceding example *cause* is a second order relation). In general, the order of a predicate is one higher than the order of its highest ordered argument; objects are always zero order (for a fuller account, see Falkenhainer, Forbus, & Gentner, 1989).

Using these definitions, the Structure-Mapping Theory of analogy holds that analogies are a special case of comparisons between two domains, where primarily the relational predicates, but not the attributes or objects, match (Gentner, 1983; see also Gentner & Markman, 1997). The two domains are called the base and the target. Usually the base domain is the one the analogizer knows more about, although it is possible that both are known equally well.

In the Structure-Mapping theory, analogical comparison and inference generation take place via a process of mapping. Mapping occurs when the predicates of the base domain are placed in correspondence with the target domain. The theory places several constraints on the mapping process including parallel connectivity and one-to-one mapping. Parallel connectivity ensures that once a correspondence between two predicates has been made, their arguments are placed in correspondence also. One-to-one mapping means that once an element of the target has been swapped out for an element of the base, that same correspondence must be honored for the rest of the mapping process. In this manner, an element in one of the domains cannot correspond to more than one element in the other domain.

Review of Analogical Inference

Analogies would be of limited use if they only found similarities between domains. Analogies can also help people extend their knowledge of a topic by virtue of its similarity to another, better known topic. An analogical

inference is a conjecture about the underlying structure of a lesser known (i.e. target) domain, based on similarities that exist between its structure and that of a better understood (i.e. base) domain. This conjecture consists of hypothesizing the existence of relations between objects (or other relations) in the target domain because those same relations are known to obtain between the corresponding objects in the base domain. Within the Structure-Mapping Theory, the inference is executed by carrying relations from the base to the target. Next, substitutions are made for matching objects, and the existence of new elements in the target may be proposed if licensed by the structure of the base domain.

The generation of inferences in this manner is widely considered to be a central aspect of analogical processing and use. Naturalistic observation of scientists engaged in the discovery process has shown that this kind of reasoning is frequently used (Dunbar, 1995). Similarly, Gentner, Brem, Ferguson, Markman, Wolff, Levidow, & Forbus (1997) found that Johannes Kepler used analogies constantly to generate ideas in the process of discovering the laws of planetary motion.

Analogical inference is arguably more important than other kinds of reasoning (e.g. syllogistic) that have received considerable scrutiny in cognitive psychology. However, very little research has been conducted on analogical inference thus far (but see Clement & Gentner, 1991; Holyoak, Novick & Melz, 1994; Krawczyk, Holyoak, & Hummel, 2005; Markman, 1997; and Spellman & Holyoak, 1996)[1].

One important issue concerning analogical inference is how the potential inferences are constrained. Once an analogy between two domains is noticed, many inferences are possible. Indeed, any fact known about the base could possibly be posited about the target, but most of these inferences would probably be false or irrelevant. This necessitates limits on the inference generation process. Clement & Gentner (1991) demonstrated that inferences are constrained by a principle of systematicity. This means that when someone is faced with a situation in which many mappings are possible, they prefer to map relations that are more deeply interconnected or that create more deeply matching systems of relations in the domains. Markman (1997) found additional support for this principle.

An issue that has not been resolved is how people evaluate inferences that are drawn from analogies. In particular, what underlying structural aspects of an analogical inference make people think that a relation is more or less likely to be true in the target domain as well? Forbus, Gentner, Everett and Wu (1997) suggested two structural aspects of analogy that might influence confidence in an inference: *support* and *extrapolation*. Figure 1 illustrates these factors. The black elements are known correspondences between the two domains, while the gray elements are known to exist in the base domain and inferred in the target. Forbus et al. (1997) define support as the amount of corresponding structure directly connected to the inference. The more support for an inference, the more confident people should be in that inference. Extrapolation is the degree to which the inference goes beyond the known correspondences. The plausibility of an inference should decrease as extrapolation increases.

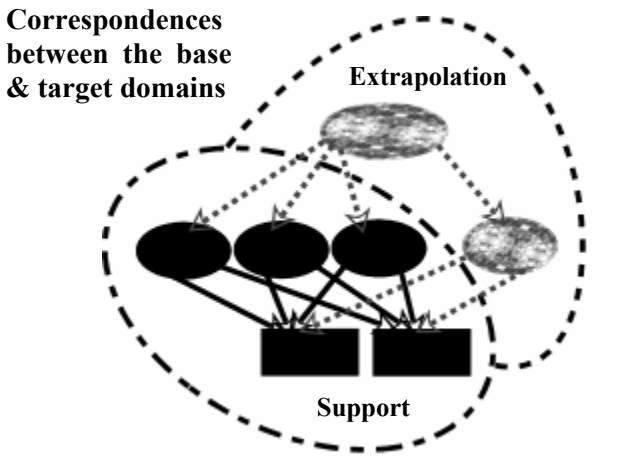


Figure 1: Structural Factors Hypothesized to Influence Evaluations of Analogical Inferences.

The Present Experiment

Our program of research examines the role of support in analogical inference.

We created a fictitious situation that allowed us to experimentally manipulate the level of support for a relation that was to be inferred in the target domain. We operationally defined support as the number of corresponding relations that were directly connected to the inference. In this study we will contrast a one relation condition with a three relation condition. If support influences inference evaluation, confidence in an inferred relation should be higher in the three relation condition. We predict that this will be the case.

Because pilot testing indicated that the materials were detailed enough to make it difficult to see how all the relationships fit together, we needed another task to ensure that the participants fully understood the materials before they evaluated the analogical inference in which we were interested. We accomplished this by having the participants evaluate the truth of a series of other propositions about the domains first. Forbus et al.'s hypotheses make no specific predictions about the subjects' responses to these statements, but these statements do allow us to examine whether the participants understood the task and were taking it seriously. We included false and unrelated propositions in order to minimize a response bias. In addition, we told the participants that we were testing a hypothesis about memory. We hoped the idea of an upcoming test would also help to motivate them to study the materials carefully.

[1] A PsychINFO keyword search for "Syllogistic Reasoning" yielded 255 hits, as opposed to the five papers listed here.

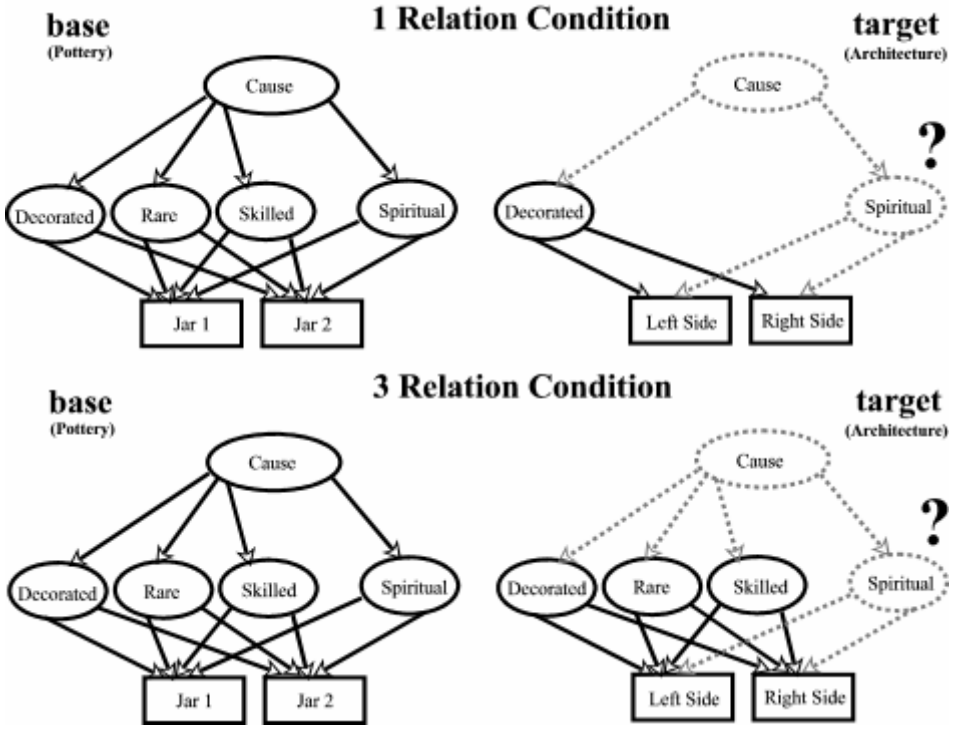


Figure 2: The structure of the domains in the one and three relation conditions.

Method

Participants

The participants were 132 undergraduates at the University of Texas at Austin (91 females, 41 males, median age = 19, range 18-29) who received course credit for their time.

Design

This experiment used a two group, between-subjects design. The main factor was Support Level. There were two Support Levels: one antecedent relation and three antecedent relations. The basic design can be seen in figure 2. The one relation condition was further subdivided into three subgroups based on which of the antecedent relations was given. There were 66 participants randomly assigned to each condition, with 22 in each of the subgroups of the one relation condition.

Materials

The materials for this study described a fictitious tribe in the West African rainforest that some cultural anthropologists want to study. The anthropologists have seen some pottery that the tribe makes, but they have never seen the tribe's houses. Nonetheless, they have heard some rumors about the houses.

We created two versions of this story. In the base domain (i.e. pottery), both versions have three antecedent relations that obtain between two kinds of jars. Collectively, these three antecedent relations cause a consequent relation to obtain. In the three relation condition, all three of these antecedent relations map onto the target domain (the construction of the tribe's houses). In the one relation version, only one of the three antecedent relations was given. Which of the three antecedent relations was given was randomly assigned.

Briefly stated, the content was as follows: In the base domain, of the two kinds of jars the tribe made, the first was more ornately decorated, made from more rare materials, and made by more skilled craftsmen (the tribe's elders), than the second kind of jars. These facts caused the first kind of jars to be considered closer to the spirit world. The target domain concerned how the tribe constructed their houses. These were one room structures that had a space to the right and a space to the left inside the door. In the three relation condition, the left area is known to be more decorated, built with rarer materials, and built by more skilled craftsmen than the area to the right. In the one relation condition, only one of these was given. The structure of the domains in the different conditions is illustrated in Figure 2[2].

[2] Note that in the figure, in the one relation condition, the decorated relation is given, while rare and skilled are not. In the actual study, there were participants who were given rare but not decorated or skilled, and participants who were given skilled but not decorated or rare as well.

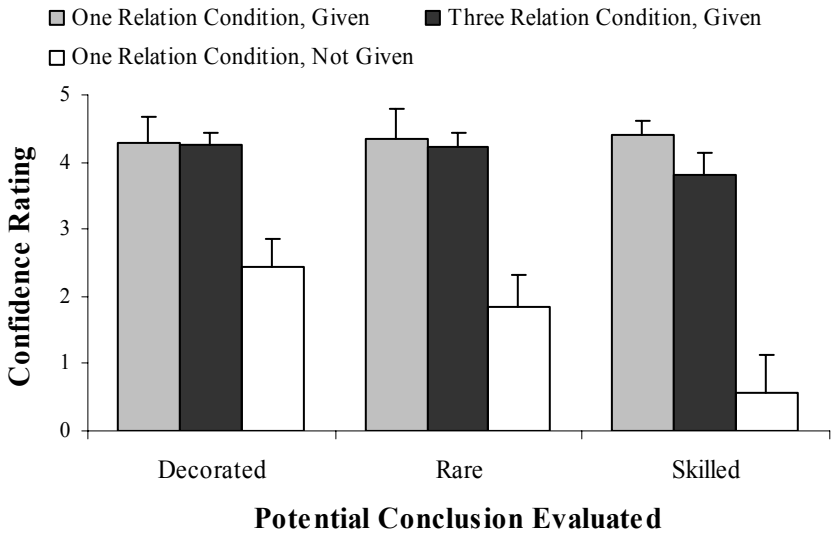


Figure 3: Mean confidence ratings for the antecedent relations. Error bars represent 1 SE.

Following the descriptions of the base and target domains were a series of seven potential conclusions that the anthropologists could draw based upon the information given. These were the same in both conditions. The potential conclusions were related to the information given in the descriptions in such a way that the participants should have known that they were certainly true or false, except for one that was completely unrelated. The false and unrelated propositions were included to minimize the potential for a response bias; they will not be discussed again. One distinction is worth noting: for participants in the one relation condition, two of the antecedent relations are not given, and thus should receive lower ratings. The potential conclusions are listed in Table 1. The first five were randomly ordered for each participant; the last two always came last in the order displayed in the table.

Other materials included a distracter packet with similarity ratings between scenes and two questionnaires (self-construal and need for cognition), and a final packet consisting of a ten item true-false memory test and demographic questionnaire. These materials allowed us to gather some additional data and facilitated the deception that the experiment was about memory, but were peripheral to the main question at issue. They will not be discussed further.

Table 1: Potential Conclusions to Evaluate

	Title	Question
1	Decorated	The space to the left of the door is much more ornately decorated than the right.
2	Rare	The materials used in the construction of the left area are more rare than those for the right.
3	Skilled	The space on the left is constructed only by the village elders.
4	False	The roofs of Brou houses are made out of bamboo.
5	Unrelated	Brou houses are built on stilts.
6	Similar	The principles and cultural practices that are involved with Brou pottery and architecture are very similar.
7	Spiritual	The area to the left of the door is considered to be closer to the spirit world.

Procedure

The participants were told that the experiment was intended to test memory for a body of material after having drawn conclusions based on it. They were told to pay very close attention to the relationships between all of the details in the first packet. The contents of the first packet are described above. After reading about the two domains, the participants rated the plausibility of the seven potential conclusions that the anthropologists could draw. First, subjects checked whether they thought the proposition was true or false. Then they had to write a few sentences justifying their response. Finally, they had to provide a confidence rating on a scale from zero to five (completely confident).[3] This setup was designed to ensure that the participants had built an accurate representation in their minds of the two domains and the correspondences between them before they made the judgment in which we were primarily interested.

Results

Seven outliers were trimmed. The overall pattern of the data did not change as a result of trimming: none of the results switched between significance and nonsignificance because of trimming.

To establish the validity of our primary analyses, we first need to show that the participants understood the task and performed rationally. This can be assessed by examining the pattern of responses to the other potential conclusions.

[3] For purposes of data analysis, the confidence ratings were scored as negative values if the participant checked that they thought the proposition was false.

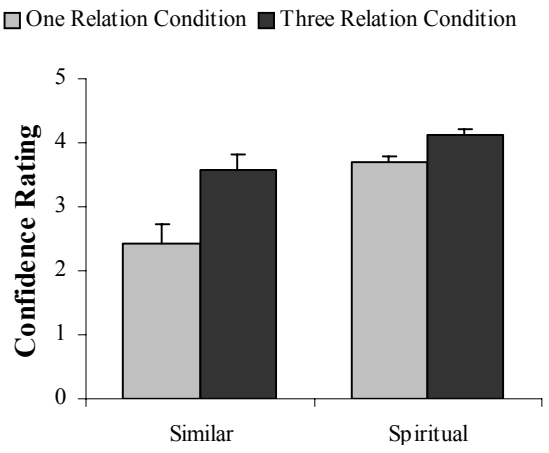


Figure 4: Mean confidence rating for the last two potential conclusions

Figure 3 shows the global pattern of results for the confidence ratings for the antecedent relations. A 2(Condition) X 3(Antecedent) ANOVA was conducted on the confidence ratings for the antecedent relations when they were given. The content of the antecedent relations was treated as a random factor for this analysis. Unsurprisingly, the test revealed no differences among the ratings (the grand mean for the antecedents when they were given was 4.22). There was no main effect of condition $F(1,2) = 2.02$, $p = 0.29$, $\eta^2 = 0.003$[4], there was no main effect for the content of the antecedent relation $F(2,2) = 0.44$, $p = 0.69$, $\eta^2 = 0.001$, and there was no interaction $F(2,247) = 0.41$, $p = 0.66$, $\eta^2 = 0.003$. These results show that the participants understood that the information they were given was solid; neither the content of the proposition nor the other information they were given influenced this.

Alternatively, when the antecedents were not given, the content of the propositions did influence the judgments that were made (decorated: $M = 2.44$, rare: $M = 1.85$, skilled: $M = 0.57$). A one-way ANOVA conducted on the antecedent relations, when they were not given, was significant $F(2,119) = 3.92$, $p = 0.02$, $\eta^2 = 0.06$. This merely implies that the content of these relations differ in their *a priori* believability.

Figure 4 displays the results for the last two potential conclusions that the participants evaluated (noted in table 1 above). Subjects in the three relation condition expressed more confidence in the proposition that the two domains were similar ($M = 3.58$) than those in the one relation condition ($M = 2.42$), $t(111^5) = 3.01$, $p = 0.003$, $d = 0.54$. Regarding the primary hypothesis of this investigation, subjects in the three relation condition ($M = 4.13$) were significantly more confident in the inference than subjects in the one relation condition were ($M = 3.69$), $t(123) = 3.33$, $p = 0.001$, $d = 0.6$.

Participants in the one relation condition could be further subdivided into three subgroups based on which of the three antecedent relations they were given. However, these subjects did not differ in their ratings for the consequent relation based on which of the three antecedent relations they had been given, $F(2,58) = 0.31$, $p = 0.73$, $\eta^2 = 0.01$. This result, coupled with those stated above, suggests that it is the amount of structural support that influences the judged strength of the inference, not the content of the supporting relations. This is true even though the antecedents differ in believability.

This experiment was designed explicitly to test the support hypothesis. However, making the three relation condition have more support than the one relation condition also caused it to be more similar. Therefore, as a post-hoc, alternative hypothesis, one could argue that support does not increase the plausibility of an inference, but rather that similarity does. Under this interpretation, increasing support would cause similarity to go up, and similarity would mediate the relationship between support and the perceived strength of an inference. To test this possibility, we ran a regression analysis predicting the ratings for the inference based on the ratings for the similarity proposition. Similarity was not a significant predictor, $b(\text{similar}) = 0.05$, $F(1,123) = 2.27$, $p = 0.13$, $r^2 = 0.02$, thus it does not seem to be driving inference evaluation.

Discussion

This study provides strong evidence for Forbus et al.'s support hypothesis. The participants were influenced by the logical structure of the information they were given. Those in the three relation condition rated the consequent relation more likely to be true, as the support hypothesis predicts.

When information is given, the propositions are evaluated without regard for their *a priori* believability, despite the fact that these propositions do vary in their innate plausibility. Furthermore, when only one antecedent relation is given, the likelihood of the consequent relation is the same without regard for the content of the antecedent relation the participants knew about. Thus, for these stories, the strength of the analogical inference is based on structural aspects, not content. These facts suggest that a judgment of an inference based on different content than that used in this study, but that had the same structural properties, would be influenced similarly. These results suggest that support is a structural factor that influences the evaluation of analogical inferences independent of content.

The results are especially impressive considering that this study was conducted using a between-groups design. The participants were not even aware that they were in a reasoning experiment, having been told that we were investigating memory. In addition, the subjects only evaluated one analogy and relevant analogical inference. Typically, in reasoning research, participants generate or evaluate several inferences and/or arguments. These within

[4] We assert the null here and several more times in the remainder of the results section. Such assertions are invalid when the investigation does not have enough statistical power to capture an effect even if it did exist. We chose to run 132 participants for this reason, and never have less than an 80% chance of resolving a large effect if there is one. Note also that in each instance where we assert the null the effect size is very close to zero.

[5] There are fewer degrees of freedom for this test because homogeneity of variance did not hold.

subjects designs boost the effect size that is found by reducing the error variance, making a more powerful test. But, this is not the only effect. They also provide the participants with an opportunity to calibrate their intuitions by exposing them to a wide range of the argument space. When there are arguments of differing strengths to be rated, the participant can become clearer about how to evaluate an argument than they would otherwise be. This does not well reflect the evaluations people make in natural situations. In addition, task demands may be introduced. It is worth reiterating that this was not the case here: we found a larger than average effect size despite the fact that participants only evaluated one inference.

Further research needs to be done in the future to develop a fuller understanding of analogical inference. Additional research is required to fully test Forbus et al.'s hypotheses as well. We are planning several follow-up experiments to replicate and extend our findings concerning *support*. Based on a qualitative examination of the justifications participants wrote for their evaluations, some individual differences dimensions (e.g. need for cognitive closure) may play a role in the evaluation of analogical inferences. In this way, we are planning to broaden our investigation. Ultimately, Forbus et al.'s extrapolation hypothesis needs to be tested as well.

Acknowledgments

We would like to thank the members of the Similarity and Cognition lab, and the Cognition and Perception area of the University of Texas at Austin, and Mariska Leunissen for helpful comments and discussion. We would also like to thank Mary Joe Amuneke and Thao Bui for help with running subjects.

References

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

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

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

Forbus, K. D., Gentner, D., Everett, J. O., & Wu, M. (1997). Towards a computational model of evaluating and using analogical inferences. *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 229-234). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Gentner, D., Brem, S., Ferguson, R., Markman, A.B., Wolff, P., Levidow, B.N., & 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., Holyoak, K. J., & Kokinov, B. N. (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.

Gentner, D. & Markman, A. B. (1997). Structural alignment in analogy and similarity. *American Psychologist*, *52*, 45-56.

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). Component processes in analogical transfer: Mapping, pattern completion, and adaptation. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory, Vol. 2: Analogical connections*. Norwood, NJ: Ablex.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping, *Psychological Review, 3*, 427-466.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models, *Cognitive Science*, *3*, 387-438.

Krawczyk, D. C., Holyoak, K. J., & Hummel, J. E. (2005). The one-to-one constraint in analogical mapping and inference. *Cognitive Science, 29*, 797-806.

Markman, A.B. (1997). Constraints on analogical inference. *Cognitive Science*, *21*, 373-418.

Spellman, B. A., & Holyoak, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology, 31*, 307-346.