Analogical Inferences in Causal Systems

Matthew Myers (matthewmyers2020@u.northwestern.edu)
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

Dedre Gentner (gentner@northwestern.edu)
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

Abstract

Analogical and causal reasoning theories both seek to explain patterns of inductive inference. Researchers have claimed that reasoning scenarios incorporating aspects of both analogical comparison and causal thinking necessitate a new model of inductive inference (Holyoak, Lee, & Lu, 2010; Lee & Holyoak, 2008). This paper takes an opposing position, arguing that features of analogical models make correct claims about inference patterns found among causal analogies, including analogies with both generative and preventative relations. Experiment 1 demonstrates that analogical inferences for these kinds of causal systems can be explained by alignment of relational structure, including higher-order relations. Experiment 2 further demonstrates that inferences strengthened by matching higher-order relations are not guided by the transfer of probabilistic information about a cause from base to target. We conclude that causal analogies behave like analogies in general—analogical mapping provides candidate inferences which can then be reasoned about in the target.

Keywords: analogy; causality; structure mapping theory; inductive inferences

Introduction

The current paper challenges recent claims that standard theories of analogy (such as structure-mapping theory; Gentner, 1983, 1989) cannot explain analogical inferences that incorporate causal relations. Holyoak and colleagues (Holyoak et al., 2010; Lee & Holyoak, 2008) contend that causal analogies require a different kind of process from typical analogies. Specifically, they claim that structure-mapping theory (SMT)—and more broadly, all extant models of analogy—fail to predict people's inferences for causal analogies that involve both generative and preventative causal relations. In their view, causal analogies require models of analogy that incorporate the basic elements of causal models (Lee & Holyoak, 2008).

We maintain that in causal analogies, the mapping between analogs is done by the same structure-mapping processes as in other domains. Assuming that the mapping yields candidate inferences in the target, normal causal reasoning processes then occur in the target to arrive at further conclusions. Specifically, we show that this division of labor holds for the kinds of materials used by Holyoak and colleagues (Holyoak et al., 2010; Lee & Holyoak, 2008): analogical processes inform the construction of causal models in the target analog, after which causal reasoning processes are used to draw further inferences in the target. We believe our account provides a better explanation of people's reasoning at the level of representation and, more broadly, offers a more parsimonious description of analogical reasoning.

Inference and Similarity

Similarity plays an important role in SMT. While simple physical or property-based similarities can serve as cues to engage in analogical comparison, relational matches are more central to the content of analogical inferences (Gentner & Markman, 1997). Relational similarity is assessed by a process of structural alignment in which components of the two analogs are placed in correspondence based on a maximal (or near-maximal) match in relational structure. Alignments with deeply embedded relational structures—in which higher-order constraining relations govern lower-order relations—are perceived as more similar than those with shallow structures (Gentner, Rattermann, & Forbus, 1993) and provide a better basis for candidate inferences to the target (Clement & Gentner, 1991). Thus, the perceived structural similarity and inferential strength between two analogs typically exhibit a positive correlation.

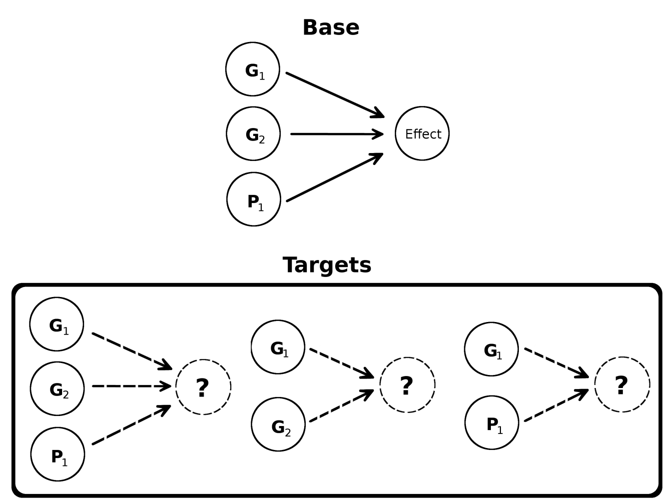

Figure 1: The causal systems in Lee & Holyoak (2008). G and P represent generative and preventative causes. The effect is the outcome feature. Dotted elements in the targets represent information not given. In descending order, similarity ratings between base and target was $G_1G_2P_1$, G_1G_2, and G_1P_1. The order of inductive strength ratings was G_1G_2, $G_1G_2P_1$, and G_1P_1.

However, Lee and Holyoak (2008) found that these measures can be disassociated. In Experiment 1, they presented participants with a description of a fictional animal with four notable features. Three of the four features were described as causally related to the fourth: two were generative causes while the other was a preventative cause (G_1, G_2, and P_1). This animal served as the base analog (see Figure 1). Participants were further given a description of a secondary target animal that possessed one of three combinations of the antecedent features described in the base (i.e., $G_1G_2P_1$, G_1G_2, and G_1P_1); however, they were given no information about how those features affected the outcome. The base animal (Animal A) and target animal (Animal B) were described in the following manner:

> Animal A has dry flaky skin, muscular forearms, a weak immune system, and blocked oil glands.
>
> For Animal A, dry flaky skin, tends to PRODUCE blocked oil glands; muscular forearms tend to PRODUCE blocked oil glands; a weak immune system tend to PREVENT blocked oil glands.
>
> Animal B has dry flaky skin, muscular forearms, and a weak immune system.

Participants' task was to either rate the similarity of the base and target, or to estimate the likelihood that the effect would occur in the target. Lee and Holyoak reported that while participants' similarity ratings roughly corresponded with the number of structural relations shared by base and target ($G_1G_2P_1$ was highest), their inferences did not. In descending order, the observed strength of the effect inference in the target was G_1G_2, $G_1G_2P_1$, and G_1P_1. The authors contend that these results are problematic for SMT and other models of analogy, for two reasons. First, SMT cannot account for the systematic non-correspondence observed between similarity and inference strength; and second, SMT cannot account for the transfer of probabilistic information from the base because it does not permit the transfer of non-relational properties of the higher-order relations[1].

In response to the first issue, Colhoun and Gentner (2009) note that the measure of inductive strength used in the previous experiment focused on a single variable: people's belief about the likelihood of the effect. This ignores a large number of other inferences that must be made between the analogs. Participants are told only that certain factors (e.g., G_1, dry flaky skin) exist in the target, but they are not told that these factors are causally connected to the effect E; these causal links must all be inferred from the base.

[1]Holyoak and colleagues (Holyoak et al., 2010; Lee & Holyoak, 2008) at points extend this claim even further, suggesting that SMT establishes analogical inferences "solely on the logical form of representations and not on their meaning" (Lee & Holyoak, 2008 p 1120). This however would suggest that SMT is insensitive to the content of the higher-order relations that bind predicates. This is incorrect. SMT distinguishes higher-order constraining relations that confer systematicity (such as *cause* and *prevent*) from non-constraining relations (such as *and*), which do not. (Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1983). Here we focus on the aspects of their argument that would prove challenging for the theory.

Arguing that the inductive strength of an analogy should be measured by all of its candidate inferences, and not solely a single effect inference, Colhoun and Gentner (2009) used Lee and Holyoak (2008) original stimuli and asked participants to rate their confidence in each of the required inferences in the target (G_1 tends to produce E, etc.). The result was that the ordering of inductive strength ratings for inferences within the target closely matched the ordering of perceived similarity. In sum, this experiment showed that when the appropriate inferential questions are asked, there is no conflict between perceived similarity and perceived inferential strength.

Inference and Structure

The second claim against SMT is that it would require the transfer of non-relational properties of higher-order relations, such as propensities for causal antecedents to produce or prevent an effect (as in G_1 tends to produce E). Lee and Holyoak (2008) claim that avoiding such a violation of systematicity would require a model to incorporate a kind of mapping process in which degrees of belief in the inferred property of a target are mutually informed by both analogical mapping procedures and causal strength assessments.

However, Colhoun and Gentner (2009) proposed an alternative solution. They suggest that the pattern of relations in the base allows people to infer that the pattern of G_1, G_2 and P_1 is sufficient to produce E. Specifically, participants are told that the base contains G_1, G_2, and P_1, that G_1 and G_2 both tend to produce E, that P_1 tends to prevent E, and that E is present. Thus in encoding the base, the presence of the effect E allows people to infer that the combined causal strength of G_1 and G_2 exceeds that of P_1. To test this, in Experiment 2, they presented participants with a base analog consisting of generative and preventive causes (G_1 tends to produce E, etc.), but varied whether the effect E was actually stated to be present in the base. Participants' task was to generate effect inferences for different targets. The result was that participants gave stronger effect inferences in the target when given a base analog in which the effect was clearly stated to be present than when given a base in which the effect might or might not be present.

Colhoun and Gentner (2009) argue that the presence of an effect in the base provides evidence of the relative strength of the antecedent causes. In other words, people are attending to and transferring a higher-order qualitative relation—that G_1 and G_2 are causally stronger than P_1. This suggests that, as in other areas of analogical reasoning, encoding processes occur in the base, followed by structure-mapping processes that align the base relational structure with that of the target and project candidate inferences to the target. Once these new inferences are projected, causal reasoning processes in the target can produce further inferences.

But there is an alternative account. Perhaps, consistent with the idea that causal propensities are intertwined with the analogical mapping process (Holyoak et al., 2010), the differences in observed inference ratings might simply have been the consequence of a probability calculation computed in the base and subsequently mapped to a matching target. For example, people may have derived a series of probabilistic propensities for the causal antecedents to bring about the effect. Participants would then project those probabilistic properties of the higher-order predicate relations onto the target (e.g., predicates related by CAUSE-10%, CAUSE-50%, CAUSE-93%, etc.). We explore this possibility in the following two studies.

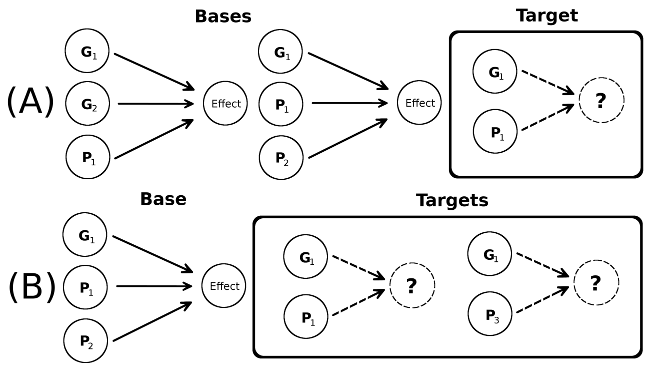

Figure 2: Base and target systems for Experiment 1 (Panel A) and Experiment 2 (Panel B). G and P represent generative and preventative causes. The experimental manipulations for Experiment 1 and 2 were the bases and targets, respectively.

Experiment 1

Hypothetically, there are two ways people may calculate probabilistic likelihoods for an effect in a base analog: a global posterior calculation, in which all causal antecedents are factored into a probabilistic estimate; and a local posterior calculation, in which only those antecedent causes that occur in the target are analyzed in the base. Experiment 1 examines a scenario where predictions for global or local posterior calculations are pitted against predictions made by standard mapping theory. Figure 2A illustrates the scenario used in this experiment. In this study, we varied the causal structure in the base, keeping the target (G_1P_1) constant in both conditions. Further, in both conditions, participants are told that the effect E occurs in the base. If participants generate a global posterior calculation for the effect in either the $G_1G_2P_1$ condition [i.e., $P(\text{Effect} \mid G_1G_2P_1)$] or the $G_1P_1P_2$ condition [$P(\text{Effect} \mid G_1P_1P_2)$], then we should see no difference in inference ratings between conditions. Since people's probability estimates are calculated conditionally on the aggregate influence of all causal antecedents, there is no unitary piece of information that can inform them as to how likely the effect will be when transferred to a G_1P_1 target[2].

But suppose instead that participants use the base to generate a local posterior calculation (considering only the relations that match with those in the target)—e.g., P(Effect — G_1P_1). In this case, they have no basis for a difference in inference ratings between $G_1G_2P_1$ and $G_1P_1P_2$. Because the strength of the individual causal antecedents is unknown, they would have no information about the degree to which the unmapped cause in the base either prevents (P_2) or contributes (G_2) to the effect. Therefore there is no reason to expect that a systematic difference in effect strength estimates between the two conditions.

In contrast, suppose participants utilize the type of encoding and structure-mapping techniques as described above. In this case, when given the $G_1P_1P_2 \to E$ base, participants should recognize that the effect of G_1 is stronger than that of both P_1P_2 and is therefore stronger than either of them alone. This relative strength relation is a higher-order relation which takes the causal relations as arguments. When this system is projected to the target (G_1P_1), participants should assume that the effect E will occur. In contrast, when given the other base, $G_1G_2P_1 \to E$, participants have no reason to infer that either G_1 or G_2 is stronger than P_1 (since generative relations outnumber preventative relations). Thus the prediction is that the effect inference will be stronger for $G_1P_1P_2$ than for $G_1G_2P_1$ bases.

Methods

Participants 40 undergraduate students from Northwestern University participated in the study for course credit. One student was excluded because of missing data points and seven were excluded for failing a comprehension check. In all, 32 participants were analyzed in this study.

Materials and Procedure The animal features and structures were those used in Experiment 1 by Lee and Holyoak (2008). Participants were given information about the two animals, a $G_1G_2P_1$ animal and a $G_1P_1P_2$ animal, which served as the different bases. They were also given a target, a G_1P_1 animal (Figure 2A). The bases were described using different sets of features, and therefore the target's features were unique to the given base. In both conditions, participants rated the likelihood of the effect in the target. Inferences were framed as suppositional queries asking participants to predict the number of animals that would exhibit the effect given 100 instances of the target. Furthermore, for the sake of completeness, similarity ratings between base and target were also obtained. These were assessed on a scale of zero to ten, with zero indicating "completely different" and ten indicating "identical". Similarity ratings always preceded inference ratings.

The experiment was conducted on a PC using the software program Qualtrics. Each participant received both $G_1G_2P_1$ and $G_1P_1P_2$ conditions. The order of the two base conditions were counterbalanced. Participants were first given the de-

[2] One alternative way participants could implement a global strategy would be to generate a posterior probability for the system of base relations, and either increase or decrease their estimate contingent upon which cause is absent in the target. For example, if participants are given $G_1P_1P_2$ as a base and G_1P_1 as a target, dropping a preventative relation (e.g., P_2) may lead them to boost their probability estimate. However, while this is certainly possible, we believe that such an explanation is far less parsimonious than an account based on higher-order relational mapping (i.e., $G_1 > P_1P_2$).

scription of a single base and the target animal. Similarity and inference question were placed directly below the description on the same page. Once participants had recorded an answer, they were prompted to continue to the next page. After leaving the page, participants were unable to review the content and answers from the previous page.

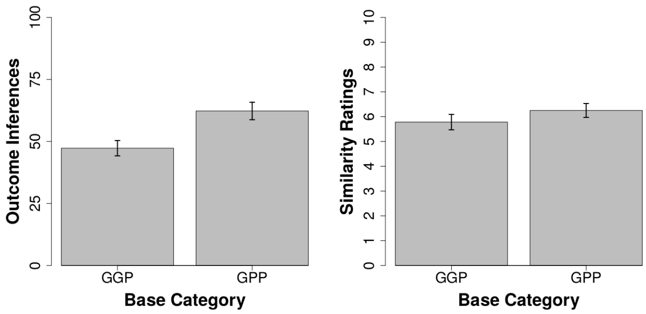

Figure 3: Mean outcome (effect) inferences and similarity ratings between target and base. The base categories represent the type of causes each cause had. Error bars represent 1 standard error of the mean.

All participants in Experiment 1 also participated in Experiment 2, which directly proceeded Experiment 1. At the end of the Experiment 2, all participants were given a multiple choice comprehension test that consisted of two questions about the content of stories they were given. If either question was answered incorrectly, that participant's data was removed from analysis for both experiments.

Results and Discussion

The mean results for both similarity and inference ratings can be seen in Figure 3. The analysis consisted of a paired-sample t-test for each rating type. There was a significant difference in mean inference ratings between the conditions: the target in the $G_1P_1P_2$ condition (M=62.28, SD=19.84) was rated as significantly more likely to exhibit the effect E compared to the target in the $G_1G_2P_1$ condition (M=47.28, SD=17.49), $t(31)=3.44$, $p<.005$. This suggests that participants' inferences were informed by the higher-order relation observed in the $G_1P_1P_2$ base. There was no significant difference in similarity ratings between the $G_1P_1P_2$ (M=6.25, SD=1.59) and the $G_1G_2P_1$ (M=5.78, SD=1.75) condition $t(31)=1.54$, p=.13. However, the relative similarity ratings were in the same direction as ratings for inferential strength.

These results show that if we assume that people inferred a higher-order relation of relative strength among the causal relations while encoding the base analog, then this relation will be mapped to the target, where it can be used to make causal inferences about the effect E. As previously discussed, this pattern of results would be highly unlikely if participants were mapping either a global or local posterior calculation from the base to target. Thus, a higher-order relationship between the generative and preventative causes in the $G_1P_1P_2$ base gave information about the relative strength of G_1 in the G_1P_1 target. This was not the case for the $G_1G_2P_1$ base. In conjunction with the findings of Colhoun and Gentner (2009), our results suggest that mapping of relational structure, including higher-order relations computed in the base, can account for inferential strength ratings made among analog causal systems.

Experiment 2

The previous study leaves open another possibility. Perhaps when given the base $G_1P_1P_2 \rightarrow E$, participants recognized that G_1 was stronger than the combined P_1P_2, but simply inferred the absolute strength of G_1. That is, they inferred that G_1 was extremely likely to produce effect E. To test this, we use the same $G_1P_1P_2 \rightarrow E$ base condition that had previously elicited increased inference ratings in Experiment 1. However, this time, the base is constant while the targets vary (see Figure 2B). In both conditions, participants receive the same generative cause (i.e., G_1) observed in the base analog, but also an additional preventative cause that differs by condition. In the Familiar P condition (G_1P_1), they receive the same P_1 feature found in the base. In contrast, in the Novel P condition (G_1P_3), they receive a novel preventative feature (i.e., P_3) that has no corresponding relation in the base. If participants simply infer extremely strong causal strength for G_1 (i.e., G_1 overpowers the preventative causes), then we should observe no difference between the two conditions. However, if participants are transferring a higher-order relational structure from the base to target, then we should find that people rate the effect to be more probable in the Familiar P condition compared to the Novel P condition.

Method

Participants The same 40 Northwestern students who participated in Experiment 1 also participated in Experiment 2. Furthermore, the same eight participants whose data was removed from analysis were likewise removed for Experiment 2. A total of 32 participants were therefore analyzed.

Materials and Procedures The animal features for Experiment 2 were taken from Colhoun and Gentner (2009). All participants were run in both the Familiar P and Novel P conditions; order was counterbalanced between participants. They were given the same similarity and inference tasks as in Experiment 1. As before, similarity queries always preceded inference ratings. Participants began the experiment immediately after finishing Experiment 1.

Results and Discussion

Figure 4 shows the mean inference and similarity ratings. As before, a paired-sample t-test was conducted for both measures. Consistent with the hypothesis, mean inference ratings for the likelihood of effect E were significantly greater in the Familiar P condition (M=62.31, SD=21.18) than in the Novel P condition (M=48.59, SD=17.59), $t(31)=3.71$, $p<.001$. Indeed, in the Novel P target, the estimates of likelihood of E

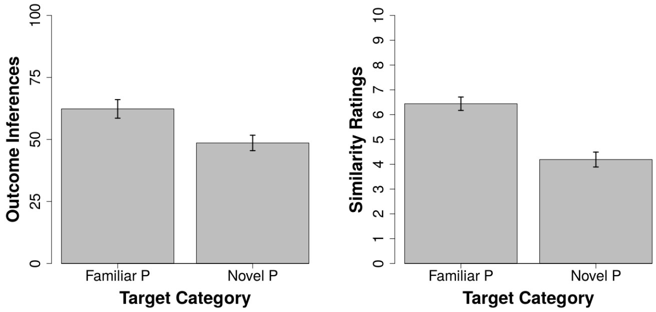

Figure 4: Mean outcome (effect) and similarity ratings between target and base. The target categories represent the type of preventative cause present in the target. Error bars represent 1 standard error of the mean.

did not differ from chance (50%). Similarity ratings were also significantly greater for the Familiar *P* condition (M=6.44, SD=1.54) than for the Novel *P* condition (M=4.19, SD=1.67), $t(31)=6.95$, $p<.001$. This is to be expected, because in the Familiar *P* condition, there are two shared factors (G_1 and P_1), while in the Novel *P* condition only one factor is shared (G_1).

These findings run contrary to the idea that participants are simply transferring absolute information about the strength of G_1 from the base to target. Had people simply transferred the strength of the generative relation, the effect inference would have been equally strong in both targets. Instead, participants only inferred that the effect occurs in the target when they could map higher-order relative strength relations from the base to the target. In sum, these findings suggest that participants' inferential strength ratings for the effect in a target can be accounted for by standard analogical mapping models.

General Discussion

Holyoak and colleagues (Holyoak et al., 2010; Lee & Holyoak, 2008) argue that causal analogies cannot be modeled in the same way as other analogies and instead require the creation of a specialized system. Specifically, they believe that most existing models, including SMT, cannot accommodate the probabilistic dynamics of causal systems. The evidence provided here suggests otherwise. Across two experiments, we demonstrate that the pattern of analogical inferences observed among various causal systems correspond with predictions made by SMT. Experiment 1 found that stronger effect inferences occurred when the causal relations in the base were united by a higher-order relation that took causes as arguments. In Experiment 2, we tested whether the results could be predicted by assuming the transfer of the individual causal strength of the generative relation from the base to target. On the contrary, the results suggest that a consistent relational structure is required in order for people to infer the effect in the target.

There has been immense progress in analogy research in the last few decades. The evidence suggests that analogy is a domain-general process that applies across physical causality (Goldwater & Gentner, 2015), mathematics (Mix, 2008; Rittle-Johnson & Star, 2007), politics (Spellman & Holyoak, 1992), spatial scenes (Doumas & Hummel, 2013; Kurtz & Gentner, 1993; Richland, Morrison, & Holyoak, 2006; Richland et al., 2006; Markman & Gentner, 1993; Sagi, Gentner, & Lovett, 2012), and scientific reasoning (Gentner, 2002; Pearl, 1992). Our findings here support the idea that analogy is a domain-general process that supports alignment and inference both within and across domains.

Acknowledgments

We thank Erin Anderson, Kensy Cooperride, Nathan Couch, Christian Hoyos, Sandy LaTourrette, Francisco Maravilla, Ruxue Shao, and Nina Simms for their comments and suggestions. This research was supported by a grant from the Office of Naval Research [award number N00014-16-1-2613] to Dedre Gentner.

References

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

Colhoun, J., & Gentner, D. (2009). Inference processes in causal analogies. In *New frontiers in analogy research: Proceedings of the second international conference on analogy* (p. 82-91). Sofia, Bulgaria: New Bulgarian University Press.

Doumas, L., & Hummel, J. (2013). Comparison and mapping facilitate relation discovery and predication. *PLOS One*, *8*, e63889.

Falkenhainer, B., Forbus, K., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*, 89-132.

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

Gentner, D. (2002). Analogy in scientific discovery: The case of johannes kepler. In L. Magnani & N. Nersessian (Eds.), *Model-based reasoning: Science, technology, values.* New York: Kluwer Academic/ Plenum Publisher.

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

Gentner, D., Rattermann, M., & Forbus, K. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, *25*, 524-575.

Goldwater, M., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition*, *137*, 137-153.

Holyoak, K., Lee, H., & Lu, H. (2010). Analogical and category-based inference: A theoretical integration with bayesian causal models. *Journal of Experimental Psychology: General*, *139*, 702–727.

Kurtz, K., & Gentner, D. (1993). Detecting anomalous features in complex stimuli: The role of structured comparison. *Journal of Experimental Psychology: Applied*, *19*, 219-232.

Lee, H., & Holyoak, K. (2008). The role of causal models in analogical inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1111-1122.

Markman, A., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, *25*, 431-467.

Mix, K. (2008). Children's equivalence judgments: Crossmapping effects. *Cognitive Development*, *23*, 191-203.

Pearl, J. (1992). *Conceptual revolutions*. Princeton, NJ: Princeton University Press.

Richland, L., Morrison, R., & Holyoak, K. (2006). Childrens development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*, 249-273.

Rittle-Johnson, B., & Star, J. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? an experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*, 561-574.

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science*, *36*, 1019-1050.

Spellman, B., & Holyoak, K. (1992). If saddam is hitler then who is george bush? analogical mapping between systems of social roles. *Journal of Personality and Social Psychology*, *62*, 913-933.