Causal Models Interact with Structure Mapping to Guide Analogical Inference

Hee Seung Lee (heeseung@ucla.edu)
Keith J. Holyoak (holyoak@lifesci.ucla.edu)
Department of Psychology, University of California, Los Angeles
Los Angeles, CA 90095 USA

Abstract

We recently proposed a theoretical integration of analogical transfer with causal learning and inference (Lee & Holyoak, 2008). A Bayesian theory of learning and inference based on causal models (Lee, Holyoak & Lu, 2009) accounts for the fact that judgments of confidence in analogical inferences are partially dissociable from measures of the quality of the mapping between source and target analogs. The integrated theory postulates a dual role for causal relations, which can guide both analogical mapping and also subsequent inferences about the target. It follows that depending on whether or not a mapping is structurally ambiguous, dropping a preventive cause from the target can either decrease or increase confidence in the same analogical inference. We report an experiment that yielded data in close agreement with predictions of the Bayesian theory. These results provide further support for the importance of integrating analogical transfer with the broader framework of causal models.

Keywords: analogical inference; causal models; mapping; Bayesian modeling.

Introduction

Analogical transfer plays a key role in scientific reasoning (Dunbar & Fugelsang, 2005). Indeed, in some areas of science in which experimental research is impossible, such as historical ethnography, analogy may provide the only viable mechanism for evaluating hypotheses. Talalay (1987) gives the example of interpreting the function of strange clay fragments discovered in Neolithic Greek sites: individual female legs, apparently never attached to torsos, that had been manufactured in pairs and later broken apart. The best clues to their function have come from other cultures in which similar tokens are known to have served to seal contracts and provide special evidence of the identity of the bearer (in feudal China, for example, a valuable piece of jade would be broken in two to mark a contract between a master and his vassal, with each keeping one piece so they could later be matched). Here the known function in a source domain has a *causal* connection to the form of relevant artifacts, and the ethnographer makes the analogical inference that a similar cause may have operated in the target domain (see Bartha, 2010).

The general question faced by a reasoner using analogy to make inferences is: Given prior knowledge at various levels of abstraction, including one or more examples that serve as source analogs, what is the probability that any potential inference about a target analog is true? For analogical inferences that involve empirical claims about the world (e.g., scientific hypotheses), answering this question depends on at least two basic subprocesses: deciding how the causally-relevant elements of the source analog(s) relate to elements of the target (structure mapping), and using the corresponding causal relations suggested for the target to estimate the probabilities of potential inferences about the target (causal inference). In the above ethnography example, mapping is required to relate the two pieces of a broken jade object to the two parts of a broken piece of pottery; causal inference is required to evaluate the probability that the clay fragments could achieve a function analogous to that achieved by a divided jade object.

Both structure mapping and causal inference have received considerable attention within cognitive science. Mapping has been the central focus of many models of analogical reasoning (e.g., Falkenhainer, Forbus & Gentner, 1989; Holyoak & Thagard, 1989; Hummel & Holyoak, 1997). Causal learning and inference have also been studied extensively, with theoretical work largely based on the framework of *causal models* (Pearl, 1988; Waldmann & Holyoak, 1992; Waldmann, Holyoak & Fratianne, 1995). The power PC theory (Cheng, 1997) provides a quantitative explanation of how the strengths of probabilistic causal relations can be learned from contingency data. More recently, this theory has been extended based on a Bayesian formulation (Griffiths & Tenenbaum, 2005; Lu et al., 2008). Theories of category-based induction have also been enriched by adopting the framework of causal models (e.g., Ahn, 1999; Kemp, Goodman & Tenenbaum, 2007; Sloman, 1994; Rehder, 2009).

Integrating analogical inference with causal models

We have proposed that a more complete understanding of analogical transfer requires specifying the role played by causal models (Lee & Holyoak, 2008; Lee, Holyoak & Lu, 2009). Figure 1 schematizes causal models for a source (left) and target analog (right). The nodes represent variable causes (C) and effects (E). The superscripts (^{S}, ^{T}) indicate the source and the target, respectively. The links represent the causal structure (only linked nodes have direct causal connections). The vectors w_i represent the causal polarity (generative or preventive) and the causal strength for links. A key assumption is that analogical transfer involves using causal knowledge of the source to develop a causal model of the target, which can in turn be used to derive a variety of inferences about the values of variables in the target. Causal relations in Bayesian causal models can carry information about existence of causal links (e.g., causal structure) and distributions of causal strength, as well as about the generating function by which multiple causes combine to influence effects.

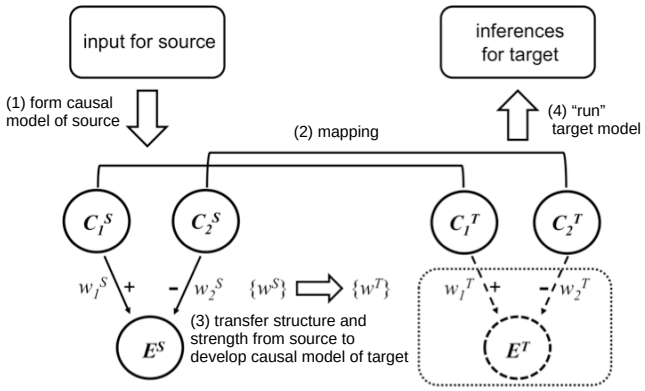

Figure 1: Framework for analogical transfer based on integrating mapping with causal models. Dotted lines indicate knowledge transferred from source to target (see text).

In our theory, the first step in analogical inference is to learn a causal model of the source. The source model is then mapped to the initial (typically impoverished) representation of the target. Based on the mapping, the causal structure and strengths associated with the source are transferred to the target, creating or extending the causal model of the latter. The model of the target can then be "run", using causal reasoning to derive inferences about the values of endogenous variables in the target. Accordingly, as summarized in Figure 1, the four basic components in analogical inference are: learning of a causal model for a source (step 1); assessment of the analogical mapping between the source and a target (step 2); transfer of causal knowledge from the source to the target based upon the analogical mapping to construct the causal model of the target (step 3); and inference based on the causal model of the target (step 4).

To generate predictive inferences (from causes to their effect), let $\boldsymbol{C}^S$ denotes the information that the source has a background generative cause, B^S, plus additional generative and preventive causal factors. (In this paper, vectors are indicated by bold font.) $\boldsymbol{C}^T$ provides analogous information about possible causes in the target. In predictive inference, the model estimates the probability of an effect occurring in the target, $E^T = 1$, based on initial information about the source, ($\boldsymbol{C}^S$, E^S), and the target, $\boldsymbol{C}^T$. The unknown causal strength of the target is represented by $\boldsymbol{w}^T$. The basic equation for predictive inference is

$$
\begin{aligned}
&P(E^T \mid \mathbf{C}^T, E^S, \mathbf{C}^S) \\
&= \sum_{\mathbf{w}^T} P(E^T, \mathbf{w}^T \mid \mathbf{C}^T, E^S, \mathbf{C}^S) \\
&= \sum_{\mathbf{w}^T} P(E^T \mid \mathbf{w}^T, \mathbf{C}^T) \sum_{\mathbf{w}^S} [P(\mathbf{w}^T \mid \mathbf{w}^S, E^S, \mathbf{C}^S, \mathbf{C}^T) P(\mathbf{w}^S \mid \mathbf{C}^S, E^S)]
\end{aligned}
$$

(Equation 1)

where the rightmost term on the right side of the equation, $P(\boldsymbol{w}^S|\boldsymbol{C}^S, E^S)$, captures the learning of a source model from observed contingency data (step 1 in Figure 1). Recent computational studies have developed detailed models that estimate distributions of causal strength by combining priors and observations (Griffiths & Tenenbaum, 2005; Lu et al., 2008). The middle term, $P(\boldsymbol{w}^T \mid \boldsymbol{w}^S, \boldsymbol{C}^S, \boldsymbol{C}^T)$, quantifies knowledge transfer based upon analogical mapping (steps 2 and 3 in Figure 1). We model the probability of transfer as

$$
\begin{cases}
P(w_i^T = w_i^S) = 1, & \text{if } C_i^T \text{ matches } C_i^S \\
P(w_i^T = w_i^S) = 0, & \text{otherwise}
\end{cases}
$$

(Equation 2)

where C_i^S and C_i^T represent the ith cause variables in the source and target, respectively. If the mapping of C_i^T to an element in the source is ambiguous (as will be the case for the materials we use in the experiment reported here), then Eq. 2 will simply sum over the transfer result obtained when C_i^T matches each of the alternative source elements, weighted by the probabilities of each of these possible matches.

The leftmost term, $P(E^T|\boldsymbol{w}^T, \boldsymbol{C}^T)$, uses knowledge from analogical transfer and observations about the presence of causal factors in the target to estimate the probability of the effect in the target (step 4 in Figure 1). For binary variables, this probability can be directly computed using the Bayesian extension of the power PC theory (Cheng, 1997; Griffiths & Tenenbaum, 2005; Lu et. al., 2008).

Mapping and Causal Inference

Although causal relations have sometimes been assumed to have special importance in guiding mapping (Holyoak, 1985; Hummel & Holyoak, 1997; Winston, 1980), models of analogical transfer have generally treated inference as a direct extension of mapping. In contrast, our causal-model approach postulates a deeper role for causal knowledge in transfer (Lee & Holyoak, 2008).

The present study sought to demonstrate a direct interaction between mapping and causal inference, which is predicted by our Bayesian theory. According to the integrated theory, a causal relation in the target potentially plays a dual role: it first may guide structure mapping between the source and target; then if mapping succeeds, it will also guide causal inference based on the resulting causal model of the target. In the present study we investigated analogical transfer when the mapping between the source and target was in some cases structurally ambiguous (cf. Spellman & Holyoak, 1996).

More specifically, we examined how presence or absence of a certain causal relation (preventive cause in this study) in the target might increase or decrease inductive strength depending on whether the structural mapping is clear or ambiguous. The source analog included a preventive cause, which might or might not be also included in the target. When the mapping is clear, the expected effect of inclusion of the preventive cause is evident in that presence of a preventive cause will decrease inductive strength in target, as shown in previous studies (e.g., Lee & Holyoak, 2008). However, when the mapping is ambiguous, and if the preventive cause is able to resolve the mapping ambiguity, the expected result will be reversed. The materials were designed so that when the mapping was ambiguous, the

inclusion of the preventive cause in the target provided sufficient structural information to resolve the ambiguity, and hence allow transfer of causal structure from source to target. Conversely, if the preventive cause were omitted from the target, the mapping ambiguity would be left unresolved, thereby impairing transfer of a causal model from source to target. In such situations our Bayesian model predicts that including the preventive cause in the target will actually *increase* inductive support for the occurrence of the effect that it actually tends to prevent. No previous analogy model predicts this type of interactive impact of causal and structural constraints on analogical transfer.

Experiment: Can a Preventive Cause Either Decrease or Increase the Judged Strength of the Same Analogical Inference?

Method

Participants Forty-five undergraduate students at the University of California, Los Angeles participated in the experiment to fulfill a course requirement. Each participant was randomly assigned to one of eight different sets of materials generated for counterbalancing purposes.

Design and materials The source story described a biochemist's findings about an imaginary liver disease called "tibulosis", found in rats. The disease had two different subtypes, "Type A" and "Type B", described as being caused by different factors and exhibiting quite different symptoms. The scientist had identified several factors that determine whether or not rats might develop either Type A or Type B tibulosis. For each type, certain hormones, enzymes, and antibodies were involved. Participants were asked to carefully study the biochemist's findings using a verbal description and diagram presented in the booklet in order to determine what characteristics are likely to produce or prevent the development of each type of the disease. Participants were then given descriptions of human patients with a liver disease, and asked to apply what they had learned about tibulosis in rats to judge the probability that the human patients had tibulosis Type A or Type B.

In the source, the two disease subtypes were designed to create a potential mapping ambiguity. The two types had identical causal structures except for the names of causal elements, but with one critical structural difference involving a preventive cause. Each source included two generative causes, one preventive cause, and an effect (consistent with a common effect model; Waldmann & Holyoak, 1992). The two generative causes were certain types of hormones and enzymes and the preventive cause was a certain type of antibody. In each case the preventive cause was narrow in scope (Carroll & Cheng, 2009), in that it served to stop the causal impact of one of the two generative causes but not the other. The description of the causal structure for Type A tibulosis was as follows:

Factors influencing development of Type A tibulosis
Hormone A tends to stimulate the production of enzyme A, and vice versa.
Hormone A tends to PRODUCE Type A tibulosis.
Enzyme A also tends to PRODUCE Type A tibulosis.
The immune system sometimes PRODUCES antibody A in response to enzyme A, but never in response to hormone A.
Antibody A tends to PREVENT enzyme A from producing Type A tibulosis. However, antibody A provides no protection against the direct effect of hormone A on Type A tibulosis.

To aid comprehension of the causal structure, a schematic diagram was provided right below the description. Figure 2 depicts the causal structure for Type A, described above. Hormone A and enzyme A are two generative causes that both tend to produce the effect, type A tibulosis. Antibody A is a preventive cause with narrow scope, which prevents enzyme A (but not Hormone A) from producing the effect. The B subtype was very similar to the A subtype described above, except that the effect was "type B tibulosis" (rather than type A), and the names of the hormone, enzyme and antibody were also B. The critical structural difference between the two sources was that in the B version, the immune system was described as producing antibody B in response to hormone B, but never in response to enzyme B (opposite to the situation in the A version); furthermore, antibody B tended to prevent the effect of hormone B (not enzyme B).

In the target story, participants read reports about human patients who might have a human form of Type A or Type B tibulosis. Examination reports for seven patients were constructed. Each examination report included information about a hormone, an enzyme, and (in some versions) an antibody found in each patient. A 2 x 2 within-subjects design was employed, resulting in four basic versions of the target descriptions. The first independent variable was whether the target description was *specific* or *generic*. In the specific condition, specific names of the hormone, enzyme, and antibody (e.g., hormone A, enzyme A, antibody A) were explicitly stated in the description of the patient report provided in the target. Given that these names matched those for one of the two subtypes described in the source,

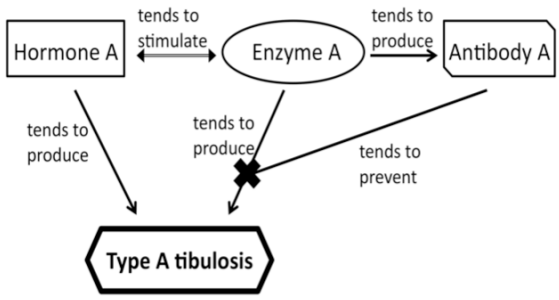

Figure 2: An example of a causal structure for one of two disease subtypes included in the source analog used in the experiment.

the mapping of the human case to Type A (or B) tibulosis as described in the source was accordingly transparent.

In contrast, in the generic condition, specific names of the hormone, enzyme, and antibody were not provided. Instead, each was simply described by its general categorical description (i.e., hormone, enzyme, and antibody). Thus in the absence of additional structural information, there was no basis for preferentially mapping the description of the factors observed in the human patient onto those related to Type A versus Type B tibulosis in rats.

The above manipulation of the target description was crossed with a second independent variable, *presence* or *absence* of the preventive cause (antibody) in the description of the human patient. Recall that the critical structural difference between Type A and Type B tibulosis as described in the source was that for Type A, the *enzyme* produced the antibody, which then acted to block the *enzyme's* impact; whereas for Type B, it was the *hormone* that produced the antibody, which then acted to block the *hormone's* impact. In the P-present condition, the target description included analogous information about the human case. For example, in the specific, P-present condition, the description might state:

> Hormone A and enzyme A are present, and each stimulates production of the other.
>
> The immune system produced antibody A in response to the enzyme (but not the hormone).

More critically, in the generic, P-present condition, the description stated:

> A hormone and an enzyme are present, and each stimulates production of the other.
>
> The immune system produced an antibody in response to the enzyme (but not the hormone).

Note that even though no specific names are provided, the above generic, P-present description (based on the second statement in the description) provides structural information sufficient to disambiguate the mapping between the human case in the target and the disease descriptions for rats as stated in the source. That is, only Type A tibulosis involves an antibody produced in response to an enzyme, which then blocked the enzyme's effect. Any of the major models of structure mapping (e.g., Falkenhainer et al., 1989; Hummel & Holyoak, 1997) would be able to use the structural information provided in the generic, P-present condition to resolve the potential ambiguity and identify a determinate mapping between the disease described in the target and one of the two subtypes described in the source. Accordingly, this condition would be essentially identical to the specific, P-present condition if participants could resolve mapping ambiguity using the preventive cause.

In the P-absent versions (both specific and generic), the second statement in the relevant description was simply replaced with "no antibody is present". Critically, in the generic, P-absent condition, no information was provided that could possibly serve to resolve the structural ambiguity inherent in the mapping; hence the target case could be mapped to either Type A or Type B in the source. If a preventive cause plays a dual role in analogical transfer, as the integrated theory postulates, then in this experiment its inclusion will have a paradoxical influence on the judged probability of an effect in the target. Specifically, given a specific description of the target, inclusion of the preventive cause will *decrease* the judged probability of the effect (by acting as a preventer within the causal model of the target); but given a generic description of the target, its presence will *increase* the judged probability of the same effect (by serving to disambiguate the mapping so that a causal model of the target can in fact be constructed).

For each condition except the generic, P-absent condition, two patient reports were constructed, resulting in seven patient reports in total. For each of the first three conditions, one of the two patient reports supported mapping to type A, and the other supported mapping to type B. Because the generic, P-absent condition did not support mapping to one type over the other, only one version of this patient report could be constructed. Two different sets of materials were constructed by counterbalancing whether the hormone or the enzyme produced an antibody in type A and in type B. Within each set, four different orders of targets were constructed, resulting in eight versions of materials in total.

Procedure Participants were given a booklet that included the source story, the target story, and a series of inference tasks. First, participants read the source story about a biochemist's findings about a new liver disease found in rats, and studied what factors were likely to produce or prevent the development of two types of the disease based on the verbal descriptions and diagrams.

In the generic conditions (but not in the specific conditions), a mapping task was included before the inference task to check if the potential mapping ambiguity was resolved or not. This task required identifying the generic hormone as "hormone A", "hormone B", or "can't tell". The analogous question was also asked about the generic enzyme. For the analogical inference task, participants were given the examination reports for seven different patients. For each, participants were asked to judge how likely it was that the patient had each disease type. To answer each question, they were to imagine there were 100 cases with the same known characteristics as for the specific case, and judge how many of these 100 cases would be expected to have each type of the disease.

Results and Discussion

On the mapping task, 33 of the 45 participants reported the structurally-justified mappings for the hormone and enzyme in the generic, P-present condition. The other 12 participants gave a variety of responses in this critical condition. We analyzed the results both including and excluding data from those participants who made mapping errors. As the basic pattern was the same in both sets of analyses, we will report the analysis including all participants.

For each patient case, participants estimated both the probability that the patient had Type A of the disease and the probability that the patient had Type B. The format

encouraged participants to treat the two types as mutually exclusive, and assignments of Type A versus Type B were fully counterbalanced across conditions. To code the responses on the inference task, we defined the "correct" disease type as that supported by the preferred mapping in the three unambiguous conditions (specific, P-present; specific, P-absent; and generic, P-present). This same disease type (either A or B) was defined as "correct" in the matched generic, P-absent condition (in which neither answer was inherently preferred).

The mean rated probability of the correct effect for each of the four conditions is shown in Figure 3 (left). These data were analyzed using a 2x2 analysis of variance in which both target description (specific vs. generic) and presence of the preventive cause (P-present vs. P-absent) were within-subjects variables. A significant main effect of specificity of the target description was obtained, $F(1, 44) = 123.09$, $MSE = 302.52$, $p < .001$, in that inference strength was significantly higher when the description was specific ($M = 83.03$, $SD = 16.09$) than when it was generic ($M = 54.26$, $SD = 17.70$). The main effect of presence of the preventive cause was not significant, $F < 1$. Most importantly, a significant interaction was obtained between target specificity and presence of preventive cause, $F(1, 44) = 79.66$, $MSE = 281.49$, $p < .001$, implying that the presence of a preventive cause had a different impact on analogical inference depending on the ambiguity of the mapping. When the description of the target was specific so that the mapping to one of the disease types in the source was transparent, participants gave significantly higher estimates of the probability of the correct effect in the P-absent condition ($M = 92.42$, $SD = 13.99$) than in the P-present condition ($M = 73.63$, $SD = 28.52$), $t(44) = 4.02$, $p = .001$. This result replicates previous findings (Lee & Holyoak, 2008; Lee et al., 2009), in that dropping a preventive cause from the target increased the strength of a predictive inference. In contrast, when the target description was generic, the effect of including the preventive cause was reversed. The estimated probability of the correct effect was now higher in the P-present condition ($M = 67.19$, $SD = 29.63$), where the preventive cause served to disambiguate the mapping, than in the P-absent condition, ($M = 41.33$, $SD = 23.89$), where the mapping was structurally indeterminate, $t(44) = 4.28$, $p < .001$.

Comparison of Bayesian Model to Human Data

We used our Bayesian model to provide a more quantitative account of our findings. The basic model was identical to that outlined by Lee et al. (2009), as summarized earlier. To fit the specific causal structures used in the present experiment, people were assumed to have no prior knowledge about causal structure or strength of the source; hence the stated causal relations were assigned a uniform strength distribution ranging between 0 and 1. Because no further information about the causal strengths was provided in the source, these distributions remained uniform (no updating based on examples), so that in effect only causal structure, not strength, was available to be transferred to the target. Based on Equation 2, causal links with uniform strength distributions were directly transferred from the source to the target analog when the mapping was determinate. Thus in the three unambiguous conditions (specific, P-present; specific, P-absent; and generic, P-present), the causal model for the correct effect was transferred to the target. In the ambiguous condition (generic, P-absent), the model summed over predictions made for each of the potential mappings to the two alternative sources, weighting them equally.

Given the general assumptions of the Bayesian version of the power PC theory (Lu et al., 2008), the predicted probability of the correct effect in the target, given the source, can be derived analytically without estimating any free parameters. To do so, the functional form of the preventive cause (a noisy-AND-NOT function) was applied in a manner that reflected the appropriate narrow scope of the preventer (Carroll & Cheng, 2009). The influences of the causes were integrated sequentially. After applying a noisy-AND-NOT function to integrate the influence of the preventer with that of its related generative cause, a noisy-OR function was applied to combine this intermediate result with the influence of the other generative cause and an assumed background cause. Figure 3 (right) depicts the parameter-free predictions of the Bayesian model. The quantitative fit was good, $r(2) = .93$. When data from just those participants who solved the mapping task correctly were modeled, the fit increased slightly, $r(2) = .94$. The model captures the trade-off that arises in the generic, P-present condition, where the presence of the preventer exerts a positive influence on analogical transfer by guiding the mapping, but then reduces transfer somewhat by acting to prevent the effect within the causal model created for the target. Also, the model makes identical predictions for the specific, P-present and generic, P-present conditions. This pattern is consistent with human response patterns in that most participants gave the same ratings for these two conditions. In the generic, P-absent condition, due to the unresolved mapping ambiguity, the probability that the effect occurs in the target is predicted by the sum of its

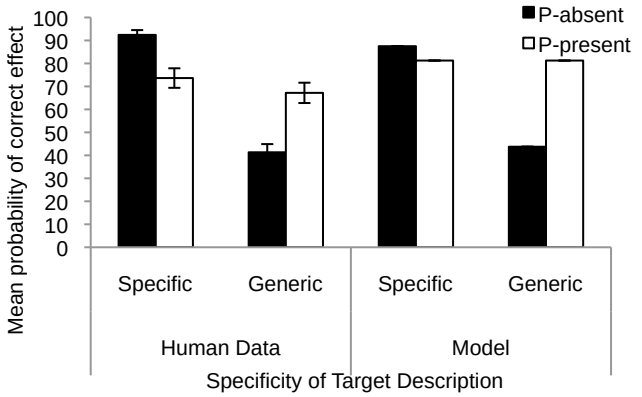

Figure 3: Mean probability of the correct effect in each condition. Left: human data; right: predictions derived from Bayesian model. Error bars represent 1 standard error of the mean.

probabilities based on each possible source, weighted by the probability of the mapping between the target and each source (equally weighted with probability of .5). Again, this prediction appears to be consistent with human response patterns, which primarily consisted of giving equal probability ratings for the two alternatives (i.e., 50/50) in this condition.

Conclusion

Our experiment demonstrated that the inclusion of a preventive cause in the target had an opposite impact on the judged probability of an effect in target, depending on whether or not the source-target mapping was ambiguous in the absence of the preventer. When the mapping was transparent (because objects in the target were described in the same specific terms as the corresponding objects in the source), inclusion of the preventive cause in the target decreased inference strength, as observed previously (Lee & Holyoak, 2008). However, when the mapping was potentially ambiguous (because objects in the target were described in generic terms), and the preventive cause provided structural information sufficient to disambiguate the mapping, then inclusion of the preventive cause in the target increased inference strength.

This pattern of interaction was predicted by our Bayesian theory (Lee et al., 2009), adding to the empirical and theoretical evidence supporting the importance of integrating theories of structure mapping with the framework provided by causal models (Waldmann & Holyoak, 1992). This type of integrated theory may provide deeper insight into many aspects of analogical inference, including its role in both the generation and evaluation of scientific hypotheses.

Acknowledgments

Preparation of this paper was supported by ONR grant N000140810186.

References

Ahn, W. (1999). Effect of causal structure on category construction. *Memory & Cognition, 27*, 1008-1023.

Bartha, P. (2010). *By parallel reasoning: The construction and evaluation of analogical arguments*. Oxford, UK: Oxford University Press.

Carroll, C. D., & Cheng, P. W. (2009). Preventative scope in causal inference. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 833-838). Austin, TX: Cognitive Science Society.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367-405.

Dunbar, K., & Fugelsang, J. (2005). Scientific thinking and reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 705-725). Cambridge, UK: Cambridge University Press.

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

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51*, 334-384.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 19 (pp. 59-87). New York: Academic Press.

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

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427-466.

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2007). Learning causal schemata. In D. S. McNamara & G. Trafton (Eds.), *Proceedings of the Twenty-ninth Annual Conference of the Cognitive Science Society* (pp. 389-394). Austin, TX: Cognitive Science Society.

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

Lee, H. S., Holyoak, K. J., & Lu, H. (2009). Integrating analogical inference with Bayesian causal models. In B. Kokinov, D. Gentner, & K. J. Holyoak (Eds.), *New frontiers in analogy research: Proceedings of the Second International Conference on Analogy* (pp. 300-309). Sofia, Bulgaria: New Bulgarian University Press.

Lu, H., Yuille, A., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115*, 955-982.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* San Mateo, CA: Morgan Kaufmann.

Rehder, B. (2009). Causal-based property generalization. *Cognitive Science, 33*, 301-343.

Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition, 52*, 1-21.

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

Talalay, L. E. (1987). Rethinking the function of clay figurine legs from Neolithic Greece: An argument by analogy. *American Journal of Archaeology, 91*, 161-169.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222-236.

Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General, 124*, 181-206.

Winston, P. (1980). Learning and reasoning by analogy. *Communications of the ACM, 23*, 689-703.