Journal of Experimental Psychology:
Learning, Memory, and Cognition
2008, Vol. 34, No. 5, 1111–1122

0278-7393/08/$12.00 DOI: 10.1037/a0012581

The Role of Causal Models in Analogical Inference

Hee Seung Lee and Keith J. Holyoak
University of California, Los Angeles

Computational models of analogy have assumed that the strength of an inductive inference about the target is based directly on similarity of the analogs and in particular on shared higher order relations. In contrast, work in philosophy of science suggests that analogical inference is also guided by causal models of the source and target. In 3 experiments, the authors explored the possibility that people may use causal models to assess the strength of analogical inferences. Experiments 1–2 showed that reducing analogical overlap by eliminating a shared causal relation (a preventive cause present in the source) from the target increased inductive strength even though it decreased similarity of the analogs. These findings were extended in Experiment 3 to cross-domain analogical inferences based on correspondences between higher order causal relations. Analogical inference appears to be mediated by building and then running a causal model. The implications of the present findings for theories of both analogy and causal inference are discussed.

Keywords: analogical inference, causal model, inductive strength, preventive cause, generative cause

In everyday life, people often face uncertainty, which permeates such diverse situations as meeting a new person, solving a novel math problem, or resolving a foreign policy crisis. To reduce uncertainty about a novel target situation, people frequently make analogical inferences on the basis of similar source situations they experienced in the past. When a source includes properties that the unfamiliar target initially lacks, any of these missing properties in the target can become a candidate for analogical inference. Yet people do not draw all possible inferences. For example, consider the earth as a source analog for inferring properties of the moon. It seems more likely the resulting analogy will lead to the inference that the moon may contain iron deposits than that the moon has a system of freeways, even though the earth contains iron deposits and also has freeways. In general, inductive inferences seem to be guided by certain general constraints that allow people to make analogical inferences selectively (Holland, Holyoak, Nisbett, & Thagard, 1986). However, the precise nature of these constraints remains to be determined.

Similarity and Structure as Inductive Constraints

There is strong evidence that people are more likely to draw confident inferences when they perceive a source and a target to be similar. According to the contrast model of similarity (Tversky, 1977), common properties tend to increase perceived similarity of two concepts whereas differences tend to reduce the perceived similarity. Many studies of category-based induction have investigated the role of similarity in analogical inference, and most of them have especially focused on similarity between premise and conclusion categories. Several studies have shown that high similarity between premise and conclusion categories promotes inference (Osherson, Smith, Wilkie, López, & Shafir, 1990; Rips, 1975; Sloman, 1993). For example, if one is told that cats have a novel biological property, this property is more likely to be ascribed to dogs than to whales. Carey (1985) showed that premise–conclusion similarity also influences inductions made by children.

Analogical inference differs from category-based induction in that analogy is generally based on single tokens (i.e., individuals) rather than categorical types. Nonetheless, similarity also plays an important role in promoting analogical transfer. By manipulating the number of shared properties of the source and target, Lassaline (1996) demonstrated that both similarity and inductive strength of analogical inferences increased with addition of shared properties. Lassaline had people either make inductive strength judgments or rate similarity on the basis of descriptions of two imaginary animals, referred to simply as *Animal A* and *Animal B*. The number of shared attributes and the number of shared relations between Animal A and Animal B were manipulated. A shared relation was defined as one possessed by both objects (source and target) but that does not connect to the attribute to be inferred in the target. An example is the following:

Animal A has X, Y, and Z.

Animal B has X, Y, and W.

For both animals, X causes Y.

Therefore, Animal B also has Z.

In this example, Animal A and Animal B have two shared attributes, X and Y, and one shared relation, X causes Y. However,

Hee Seung Lee and Keith J. Holyoak, Department of Psychology, University of California, Los Angeles.

Preparation of this article was supported by ONR Grant N000140810186 to Keith Holyoak. Preliminary reports of portions of this work were presented at the 29th (Nashville, TN, August 2007) and 30th (Washington, DC, July 2008) Annual Conferences of the Cognitive Science Society. We thank Paul Bartha and Patricia Cheng for helpful discussions. Miriam Bassok and Kevin Dunbar provided valuable comments on a draft.

Correspondence concerning this article should be addressed to Hee Seung Lee, Department of Psychology, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90095-1563. E-mail: heeseung@ucla.edu

Attribute X is not causally connected to the target property to be inferred, Attribute Z.

The results showed that the more attributes were shared between two animals, the stronger were inductive judgments as well as the higher similarity was perceived to be. However, the number of shared relations such as X causes Y (i.e., causal links that did not relate to the target property) influenced only similarity ratings: Similarity ratings increased with the addition of such a shared relation whereas inductive strength judgments did not increase. Lassaline's (1996) findings thus demonstrated that similarity influences but does not solely determine the strength of analogical inferences. (For evidence from studies of category-based induction supporting a similar conclusion, see Rips, 1989; Smith & Sloman, 1994.)

Computational models of analogy have placed major emphasis on the role of structural parallels between relations in the source and target. The importance of formal structure provided the basis for Gentner's (1983) structure mapping theory, which has been implemented in the structure mapping engine (SME; Falkenhainer, Forbus, & Gentner, 1989). Gentner distinguished between first-order relations, which take objects as arguments (e.g., "The dog is bigger than the cat") and higher order relations, which include propositions as arguments (e.g., "Because the dog is bigger than the cat, the cat ran away from the dog"). Gentner argued that higher order relations (which notably include cause as a special case) are more important for analogical inference than are first-order relations. The priority of higher order relations is due to what she termed the *systematicity* principle, which hypothesizes a preference for inferences based on predicates having many mutually interconnecting relationships. In SME, a structurally consistent set of potential mappings and candidate inferences (termed a *Gmap*) receives a *structural evaluation score*, which increases with the number of matching higher order relations. Those candidate inferences associated with the Gmap receiving the highest evaluation score are transferred to the target analog. SME thus predicts that analogical inferences will be stronger when supported by a greater number of matches between higher order relations.

In addition to SME, other computational models of analogical inference, such as the analogical constraint mapping engine (ACME; Holyoak & Thagard, 1989) and learning and inference with schemas and analogies (LISA; Hummel & Holyoak, 1997, 2003) incorporate similar relation-based constraints. Like SME, these models generate candidate inferences using variants of a procedure termed copy with substitution and generation (CWSG; Holyoak, Novick, & Melz, 1994), in which inferences about the target are constructed directly from the mapping between source and target relations by enforcing structural consistency. In all of these models, matching higher order relations (including cause) increase the evidence for analogical correspondences and hence indirectly support inferences.

Causal Knowledge as a Basic Constraint on Analogical Inference

A central issue in understanding constraints on analogical inference is whether causal relations have some special status by virtue of the fact that causes actually produce (or sometimes prevent) their effects or whether causal relations simply operate as special cases of higher order relations, defined by the formal properties of predicate-argument structures. At one extreme, the systematicity principle explicitly eschews any role for the meaning of relations in guiding analogical reasoning: "The processing mechanism that selects the initial candidate set of predicates to map attends only to the structure of the knowledge representations for the two analogs, and not to the content" (Gentner, 1983, p. 165). As noted above, in SME, the structural evaluation scores for Gmaps are based on mappings that include potential inferences computed by a CWSG algorithm; thus inferences about the target, as well as mappings between preexisting information about the two analogs, are directly determined by the systematicity principle.

In contrast, some theorists in philosophy of science (Hesse, 1966; Bartha, in press), artificial intelligence (Winston, 1980), and psychology (Holyoak, 1985) have argued that analogical inference is specifically constrained by causal understanding of the source and target. Holyoak (1985) emphasized that causal knowledge is the basis for pragmatic constraints on analogical inference rather than simply structural constraints. He noted that in the case of problem solving by analogy, "the goal is a *reason* for the solution plan; the resources *enable* it; the constraints *prevent* alternative plans; and the outcome is the *result* of executing the solution plan" (Holyoak, 1985, p. 70). Once it is recognized that the ultimate goal of analogical inference is to predict the presence or absence of some outcome in the target, it is clear that a critical consideration in analogical inference is how each factor influences the outcome in the source domain. Causal relations play a central role in making these determinations.

A number of studies have provided evidence for an important role of causal relations in category-based induction. For example, Sloman (1994, 1997) demonstrated that people are more willing to accept an inductive conclusion when the premise and conclusion categories share a causal explanation. He manipulated the presence versus absence of shared causal explanations between premises and conclusions using realistic materials such as the following:

1. Many ex-cons are hired as bodyguards.

Therefore, many war veterans are hired as bodyguards.

2. Many ex-cons are unemployed.

Therefore, many war veterans are unemployed.

The premise and conclusion in Situation 1 share a common causal explanation (both ex-cons and war veterans are experienced fighters), whereas those in Situation 2 do not share any obvious causal explanation. Sloman found that people were more confident of the conclusion for arguments such as Situation 1 than for those such as Situation 2.

In another study of category-based generalizations, Rehder (2006) showed that such generalizations are preferentially guided by causal relations, such that standard effects of typicality, diversity, and similarity itself are eliminated when a causal relation is present. However, in Rehder's experiments, the single causal relation, when present, was also the sole higher order relation. Given this inherent confounding, Rehder's findings (as well as similar results reported by Read, 1984) can be interpreted as evidence for the importance of higher order relations rather than of causal relations per se and hence are broadly consistent with Gentner's (1983) systematicity principle. However, other experimental evidence supports the general claim that category-based inductive inferences are based on causal models (e.g., Ahn, 1999; Lien & Cheng, 2000; Rehder & Burnett, 2005; see Rehder, 2007). In addition, recent theoretical work on category-based generalization has focused on the importance of causal knowledge (Tenenbaum, Kemp, & Shafto, 2007).

Experimental evidence also suggests that causal relations guide analogical transfer. Using complex stories, Spellman and Holyoak (1996) showed that when the source–target mapping was ambiguous by structural criteria, those relations causally relevant to the reasoner's goal determined the preferred mapping and inferences about the target. Using the imaginary-animal materials described above, Lassaline (1996) demonstrated that when a causal relation in the source is unmapped and the causal property is shared by the source and target, then people are likely to infer the corresponding effect in the target. For example:

> Animal A has properties X, W, and Z.
>
> For Animal A, X causes Z.
>
> Animal B has X, W, and Y.
>
> Therefore, Animal B also has Z.

Here Property X is the causal property shared by Animal A and Animal B, leading to the inference that Effect Z found in Animal A will also be present in the target, Animal B.

Moreover, Lassaline (1996) also demonstrated that people make stronger inferences on the basis of the higher order relation, "cause," than on the basis of a noncausal relation, "temporally prior to," which appears to have the same formal structure. Thus the same syntactic order of relations does not always yield the same degree of inductive strength about the target property to be inferred. Lassaline's findings have been simulated using the LISA model (Hummel & Holyoak, 2003), which assigns greater attentional resources to causal relations during its sequential mapping process.

Causal Models as Inference Engines

Although some computational models of analogical inference postulate a special role for causal relations in guiding analogical processing (Hummel & Holyoak, 1997, 2003), models of analogy have not been closely connected to models of human causal reasoning. In the present article, we explore the possibility that people may use causal models to guide analogical inference. Graphical representations of causal links have been used extensively in work on causal reasoning in philosophy (Reichenbach, 1956; Salmon, 1984), artificial intelligence (Pearl, 1988), and psychology (Cheng, 1997; Griffiths & Tenenbaum, 2005; Waldmann & Holyoak, 1992). Causal models postulate that causes can be either *generative* (making the effect happen) or *preventive* (stopping the effect from happening; see Cheng, 1997; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2006, in press). A generative cause increases the probability of an outcome whereas a preventive cause decreases the probability of the outcome. Because generative and preventive causes exert their power in opposite directions, the distinction between generative and preventive causes is crucial in predicting the outcome.

In philosophy of science, Bartha (in press) has recently extended Hesse's (1966) work on the role of causal models in analogy. Bartha distinguished between *contributing* causes (generative) and *counteracting* causes (preventive) in assessing the normative strength of arguments by analogy. He pointed out that the absence of a correspondence in the target for a counteracting cause present in the source might actually strengthen an argument from analogy. For example, iron deposits are still present on earth despite the fact that humans have been extracting iron ore in mining operations for centuries. Taking the earth as a source analog for the moon, the fact that no mining operations have so far been conducted on the moon (a mismatch with a property of the source) seems to strengthen the analogical inference that iron deposits remain to be found on the moon.

Figure 1 shows how people might reach different inductive conclusions about the probability of a possible target property on the basis of the presence or absence of a preventive cause in the target. The source has four properties: G_1, G_2, P_1, and E. Properties G_1 and G_2 are generative causes that increase the probability of Outcome E occurring, whereas Property P_1 is a preventive cause that decreases the probability of Outcome E occurring. Outcome E thus occurs in the source despite the countervailing influence of P_1. Target 1 has three properties, G_1, G_2, and P_1, all shared with the source, whereas Target 2 has only two properties, G_1 and G_2 (i.e., the preventive property P_1 is absent in Target 2). Given the same source, which of Target 1 and Target 2 will yield a stronger analogical inference about the presence of Outcome E?

All extant computational models of analogical inference, including SME (Falkenhainer et al., 1989), ACME (Holyoak & Thagard, 1989), and LISA (Hummel & Holyoak, 2003), predict that people will draw a stronger analogical inference about Target 1 than Target 2, because Target 1 shares more properties with the source than does Target 2. In particular, if one assumes that the causal connections represented by arrows constitute higher order relations, Target 1 shares three higher order relations with the source, whereas Target 2 shares only two higher order relations. Because both similarity and structural constraints on mapping solely concern correspondences of properties and relations between the source and target, models based on only these constraints predict that Target 1, which has more correspondences between the source

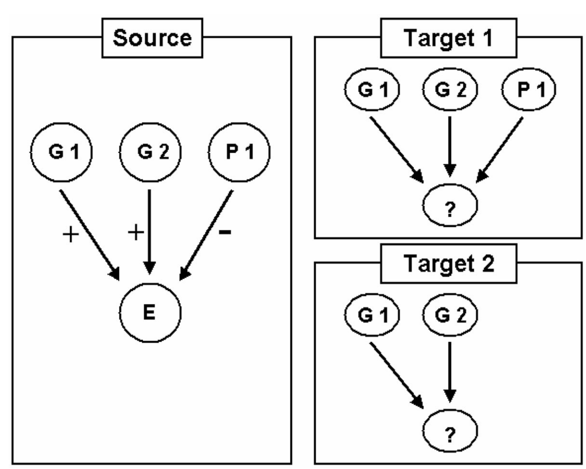

Figure 1. Example of use of causal models in analogical inference. G, P, and E represent a generative cause, a preventive cause, and an effect, respectively.

and target, will yield a stronger analogical inference than Target 2, which has fewer.

However, if people use causal models in analogical inference, then Target 2 will actually yield greater inductive strength than Target 1. Although Target 1 is more similar to the source than is Target 2, Target 1 includes the preventive cause, P_1, and this preventive cause will decrease the probability of Outcome E. In contrast, even though Target 2 is less similar to the source than is Target 1, because it includes only generative causes, G_1 and G_2, and not the preventive cause, P_1, the probability of Outcome E will be increased. Under the causal-model view, a CWSG algorithm may propose Outcome E; however, the strength of the inference will be determined not simply by CWSG alone but also by running the resulting causal model of the target to assess the probability that the causal factors would generate (or prevent) the proposed outcome. We performed three experiments to test these competing predictions.

Experiment 1

In Experiment 1, we investigated the effect of a preventive causal relation on analogical inference, using a paradigm adapted from that of Lassaline (1996).

Method

Participants. Sixty undergraduates at the University of California, Los Angeles (UCLA), received course credit for participating in the experiment. Half of the participants provided inductive strength judgments, and the other half provided similarity ratings.

Design and materials. Participants read a description of two imaginary animals referred to as *Animal A* and *Animal B,* and then they evaluated either the inductive strength of an analogical inference or the similarity of the two animals. The source analog (Animal A) always had three causal properties related to one effect property, E, consistent with a common-effect model (Waldmann & Holyoak, 1992). Across the arguments, the number of shared properties and the presence or absence of a preventive relation connecting one shared property to a nonshared property were manipulated. There were two independent variables. The first independent variable was the presence versus absence of the preventive property. In the *generative-only* condition, the source did not have the preventive property but instead had three generative properties. In the *generative + preventive* condition, the source had two generative properties and one preventive property. Each generative property tended to produce E whereas the preventive property tended to prevent E. People appear to have prior beliefs that genuine causes have high strength (Lu et al., 2006, in press) and by default may interpret causes as deterministic. However, in general, causes can be probabilistic (Cheng, 1997). Accordingly, the phrase "tends to" was included in each causal premise to make it clear that all causes were probabilistic.[1]

The second independent variable was argument type. The generative + preventive condition included three argument types: G_1G_2P, G_1G_2, and G_1P, where G and P represent a generative cause and a preventive cause, respectively. In the generative-only condition, because there was no preventive property, only two argument types were possible: $G_1G_2G_3$ and G_1G_2 (counterbalanced with G_1G_3). An example of argument type G_1G_2P is the following:

> 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 tends to PREVENT blocked oil glands.
>
> Animal B has dry flaky skin, muscular forearms, and a weak immune system.
>
> Therefore, Animal B also has blocked oil glands.

For similarity ratings, the same argument lists were used, but each argument included only the premises without the conclusion sentence.

A causal-model analysis predicts that in the generative-only condition, similarity ratings and inductive strength judgments will follow the same pattern: The $G_1G_2G_3$ argument will have higher perceived similarity and higher inductive strength than the G_1G_2 argument. However, in the generative + preventive condition, similarity ratings and inductive strength judgments will follow different patterns. The G_1G_2P argument will have higher perceived similarity than the G_1G_2 and G_1P arguments, but the G_1G_2 argument will have higher inductive strength than the G_1G_2P or G_1P arguments. In addition, the G_1G_2P argument will have higher inductive strength than the G_1P argument. In contrast, all extant computational models of analogy predict that similarity and inductive strength will be positively correlated regardless of the content of the causal relations.

Six property lists were created and six sets were constructed by counterbalancing which property list was assigned to each condition and argument type. Each participant (in both inference and similarity-rating conditions) was given a booklet consisting of six descriptions of animal pairs, referred to as *Animal A* and *Animal B*. Three of the six items were $G_1G_2G_3$, G_1G_2, and G_1G_3 arguments (generative-only condition), and the other three items were G_1G_2P, G_1G_2, and G_1P arguments (generative + preventive condition). Within each set, the order of items was randomized for each participant.

Procedure. Participants were tested individually in a small testing room. Instructions and experimental trials were self-paced and there was no time limit. Both groups of participants (similarity rating and inductive strength judgment groups) were given a booklet that included instructions and six arguments. Participants were instructed that they were to assume all of the information given in the descriptions was true. Each participant judged either how likely a conclusion would be true or how similar the pairs of animals were on the basis of the information given in the description.

For the group making inductive strength judgments, the task after reading descriptions of Animal A and Animal B (the premise statements) was to judge how likely it was that Animal B has a certain property (the conclusion statement). These judgments were solicited using a counterfactual frequency-based scale adapted from the measure of causal strength introduced by Buehner, Cheng, and Clifford (2003). In making their judgments, participants were asked to imagine there were 100 examples of Animal B and to estimate how many out of these 100 cases would have the

[1] Results similar to those observed in the present Experiment 1 were obtained in an experiment in which the phrase "tends to" was omitted (Lee & Holyoak, 2007, Experiment 1).

property stated as the conclusion, assigning a number between 0 and 100 for each item.

For the group making similarity ratings, participants were given only premise statements with descriptions of the two animals, not a conclusion statement. They evaluated how similar Animal A and Animal B were on the basis of the descriptions they read. For each description of two animals, a similarity rating scale from 0 to 10 was provided. Under the numbers 0 and 10, the words *totally different* and *identical* were written, respectively. Participants were asked to try to use the entire scale but to feel free to use any number as often as they felt it was appropriate.

Results and Discussion

Similarity ratings and inductive strength judgments were analyzed separately. For the generative-only condition, the G_1G_2 and G_1G_3 argument types were literally the same (differing only by counterbalancing), so these data were collapsed together for both similarity ratings and inductive strength judgments. The results of similarity ratings are shown in Figure 2 (top). In the generative-only condition, the mean similarity ratings for analogies based on the $G_1G_2G_3$ and G_1G_2 argument types were 8.17 and 6.62, respectively. These mean similarity ratings for $G_1G_2G_3$ and G_1G_2 arguments were reliably different, $t(29) = 4.44$, $p < .001$, such that perceived similarity increased by 1.55 points from two shared attributes to three shared attributes.

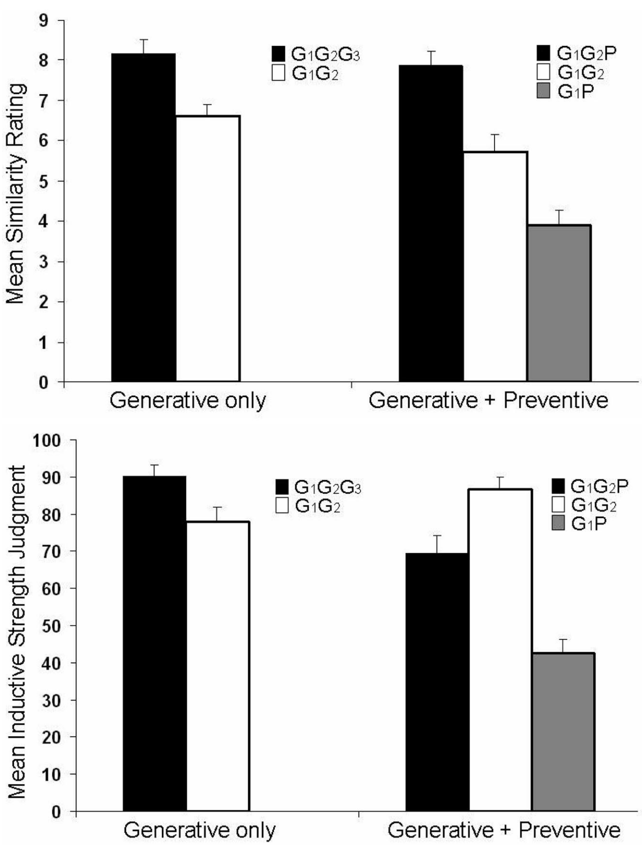

Figure 2. Mean similarity ratings (top) and mean inductive strength judgments (bottom) for each argument type in the generative-only and generative + preventive conditions of Experiment 1. Error bars represent 1 standard error of the mean.

In the generative + preventive condition, the mean similarity ratings showed a pattern similar to that observed in the generative-only condition. The mean similarity ratings for G_1G_2P, G_1G_2, and G_1P argument types were 7.87, 5.73, and 3.90, respectively. A one-way analysis of variance (ANOVA) was performed to examine the differences among the three argument types, G_1G_2P, G_1G_2, and G_1P. This ANOVA revealed a significant effect of argument type, $F(2, 58) = 29.79$, $MSE = 3.97$, $p < .001$, such that perceived similarity ratings increased from two shared properties to three shared properties. Also, even though G_1G_2 and G_1P arguments have the same number of shared properties (two), G_1G_2 arguments were rated as having higher similarity than G_1P arguments, $t(29) = 3.92$, $p < .001$.

The results for inductive strength judgments are shown in Figure 2 (bottom). For inductive strength judgments, the generative-only condition and the generative + preventive condition showed different patterns. In the generative-only condition, the mean inductive strength judgments for $G_1G_2G_3$ and G_1G_2 arguments differed, $t(29) = 3.29$, $p < .01$, such that inductive strength increased by 12 points from two shared attributes to three shared attributes. The mean inductive strength judgments of $G_1G_2G_3$ and G_1G_2 arguments were 90.27 and 78.25, respectively.

However, in the generative + preventive condition, the results for inductive strength judgments showed a different pattern. The mean inductive strength judgments for G_1G_2P, G_1G_2, and G_1P arguments were 69.4, 86.6, and 42.5, respectively. A one-way ANOVA was performed to examine the differences among the three argument types, G_1G_2P, G_1G_2, and G_1P. This analysis revealed a significant effect of argument type, $F(2, 58) = 37.03$, $MSE = 400.35$, $p < .001$. Unlike the pattern for similarity ratings, Argument G_1G_2 was rated as having higher inductive strength than either Argument G_1G_2P, $t(29) = 4.87$, $p < .001$, or Argument G_1P, $t(29) = 8.13$, $p < .001$. Also, Argument G_1G_2P was rated as having higher inductive strength than Argument G_1P, $t(29) = 3.84$, $p < .01$.

To summarize, in Experiment 1, similarity ratings generally increased with the number of shared properties, but inductive strength judgments were reduced by the presence of a shared preventive property. Similarity ratings increased with the addition of shared attributes between the source and target in both the generative-only and the generative + preventive conditions. In the generative + preventive condition, however, analogies based on argument type G_1G_2 were rated as having higher similarity than were those based on argument type G_1P, even though the number of shared attributes was the same. One possible explanation of this difference is that people may have sometimes made use of a causal model in making similarity comparisons. In Argument G_1G_2, because there are only generative factors, people may have considered Effect E to be probable and therefore inferred that the target would actually share three properties with the source: G_1, G_2, and inferred outcome E. In contrast, because Argument G_1P includes a preventive property, people may have considered the probability of Effect E to be low, thus inferring that the target would have only two shared attributes: G_1 and P.

The results for inductive strength judgments showed a different pattern from those of similarity ratings only in the generative + preventive condition, not in the generative-only condition. In the generative-only condition, Argument $G_1G_2G_3$ had higher inductive strength than did Argument G_1G_2, the same pattern as for similarity ratings. This result confirms that people consider the number of generative causes when inferring the presence of the outcome in the target. However, in the generative + preventive condition, Argument G_1G_2P had lower inductive strength than Argument G_1G_2. If people just focused on the number of correspondences between the source and the target, as the similarity and structural views assume, Argument G_1G_2P should have yielded higher inductive strength than Argument G_1G_2. Finally (and not surprisingly), Argument G_1G_2P had higher inductive strength than Argument G_1P. Thus even when a preventive cause is present, people also consider the number of generative causes.

Overall, the results of Experiment 1 suggest that people use causal models to guide analogical inferences and do not base their inferences purely on formal correspondences between the two analogs. People appear to consider not only the sheer presence of correspondences between the source and target but also whether the shared causal relations are generative or preventive.

Experiment 2

Experiment 2 was designed to investigate whether people also use causal models during analogical inference when the source includes more complex systems of higher order relations. In Experiment 1, one causal property was always connected to one effect property attributed to the same entity. For example, Animal A might have Properties X and Y, with X causing Y. In this type of case, the causally related properties (X and Y) can be viewed as features of a single entity (Animal A). Thus although causal relations are often described as higher order relations, it could be argued that the particular causal relations used in Experiment 1 held between simple attributes of a single entity rather than propositions and hence were actually first-order relations as defined by Gentner (1983).

To ensure that our basic results continue to hold when the causal relations in our inference task are clearly higher order, we developed a new set of materials in which cause–effect relations held between relations connecting multiple objects. Schematically, the form of the causal relations we used in Experiment 2 was [A is greater than B] causes [E]. In this example, there are three objects, A, B, and E. Objects A and B enter into a first-order relation (i.e., *greater than*). This first-order relation between Objects A and B is causally connected to Effect E. This structure of the causal relations clearly meets Gentner's (1983) definition of a higher order relation, as the causal relation takes a proposition (e.g., *A is greater than B*) as an argument. By investigating whether people still use causal models in evaluating analogical inferences on the basis of this more complex type of structure, it will be possible to verify the general role of causal models as a major inductive constraint on analogical inference.

Method

Participants. Sixty undergraduate UCLA students received course credit for participating in the experiment. Half of the participants provided inductive strength judgments, and the other half provided similarity ratings.

Materials and procedure. Participants read a description of a fanciful "newly discovered species of bird," which specified three causal relations between the relative amount of hormones, enzymes, and neurotransmitters (causes) and an abnormal characteristic (effect) found in the bird. The relative amount of two substances of each of two types was described as tending to produce the abnormal characteristic, and the relative amount of two substances of the third type was described as tending to prevent the abnormal characteristic. An example description is the following:

> [Hormone A > Hormone B] tends to PRODUCE blocked oil glands.
>
> [Neurotransmitter X > Neurotransmitter Z] tends to PRODUCE blocked oil glands.
>
> [Enzyme P > Enzyme Q] tends to PREVENT blocked oil glands.

After reading the description of findings for the first described bird, participants studied reports about the relative amounts of hormones, neurotransmitters, and enzymes found in a second bird of the same species. Four different types of arguments were created: G_1G_2X, G_1G_2P, G_1G_2, and G_1P. As in Experiment 1, G and P represent generative and preventive causes, respectively. The factor X was constructed by switching the order of objects in the P relation; for example, in the above example, X was [Enzyme Q > Enzyme P]. In Experiment 2, unlike Experiment 1, both G and P represent first-order relations (*greater than*) between two objects rather than properties of a single object.

Eight different descriptions of birds were created, and each argument type was assigned to each description, creating 32 items altogether. Of the total of 32 items available, 8 items were used to create a booklet for each participant, 2 of each argument type (G_1G_2X, G_1G_2P, G_1G_2, and G_1P). This counterbalancing generated four different sets of materials, thereby avoiding repeated use of the same abnormal characteristic for an individual participant. Within each set, the order of items was randomized for each participant.

The basic procedures were the same as in Experiment 1. For the group making similarity judgments, the task was to rate how similar the two birds are. For the group making inductive strength judgments, the task was to judge how likely it was that a second bird would have the abnormal characteristics described in the first bird on the basis of the descriptions of the relative amounts of hormones, neurotransmitters, and enzymes found in the two birds. The scales used for each task were of the same basic form as those used in Experiment 1.

Results and Discussion

The results for both similarity ratings and inductive strength judgments are shown in Figure 3. Similarity ratings and inductive strength judgments were analyzed separately. For each dependent measure, a one-way ANOVA was performed with the four argument types (G_1G_2X, G_1G_2P, G_1G_2, and G_1P) as a within-subjects variable.

The mean similarity ratings of analogies based on G_1G_2X, G_1G_2P, G_1G_2, and G_1P arguments were 6.35, 8.73, 7.57, and 5.03, respectively. These means were significantly different, $F(3, 87) = 24.55$, $MSE = 3.09$, $p < .001$. G_1G_2P arguments were rated as having the highest similarity of the four argument types. G_1G_2P arguments were rated as having higher similarity than those arguments with fewer shared relations, either those excluding P (mean of G_1G_2X and G_1G_2 arguments), $t(29) = 3.91$, $p = .001$, or those excluding G (G_1P arguments), $t(29) = 8.50$, $p < .001$. In addition, G_1G_2 was rated as having higher similarity than G_1G_2X (where X

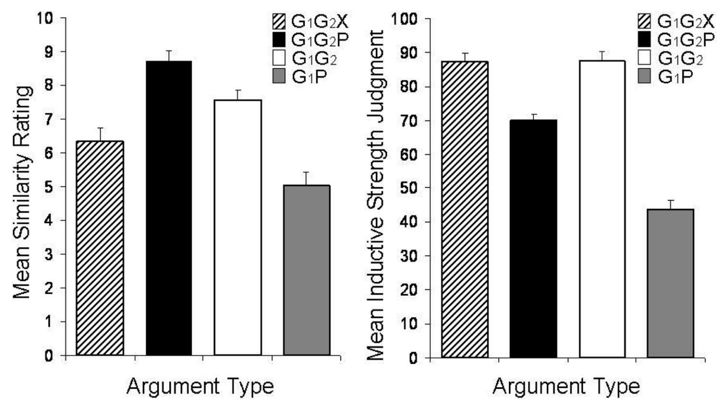

Figure 3. Mean similarity ratings (left) and mean inductive strength judgments (right) for each argument type in Experiment 2. Error bars represent 1 standard error of the mean.

was formed by reversing the relation P in the source), $t(29) = 3.84$, $p = .001$.

The mean inductive strength judgments for G_1G_2X, G_1G_2P, G_1G_2, and G_1P arguments were 87.40, 69.83, 87.47, and 43.70, respectively. The mean inductive strength judgments differed across argument types, $F(3, 87) = 75.17$, $MSE = 170.53$, $p < .001$, but showed a pattern different from that obtained for similarity ratings. G_1G_2P was evaluated as having lower inductive strength than the arguments that dropped preventive relations (mean of G_1G_2X and G_1G_2 arguments), $t(29) = 5.78$, $p < .001$, but as having higher inductive strength than the arguments that dropped generative relations (G_1P arguments), $t(29) = 8.12$, $p < .001$. Unlike the pattern for similarity ratings, inductive strength judgments did not differ for G_1G_2X and G_1G_2 arguments, $t(29) = .02$, $p = .98$.

To summarize, the results of Experiment 2 replicated the basic findings of Experiment 1 using more complex causal structures that ensured the causal relations were higher order, as defined by Gentner (1983). First, similarity ratings increased with the addition of shared relations, regardless of whether the relation was generative or preventive. G_1G_2P arguments had three shared relations between the source and target, and analogies based on these arguments were evaluated as having higher similarity than those based on arguments that had two shared relations (G_1G_2X, G_1G_2, and G_1P). It should be noted that Argument G_1G_2X was rated as having lower similarity than Argument G_1G_2. Even though Arguments G_1G_2X and G_1G_2 had the same number of shared relations between the source and target (i.e., two shared relations, G_1 and G_2), X was not present in the source, so this relation constituted a difference between the source and target, reducing the perceived similarity. This result is consistent with the contrast model of similarity (Tversky, 1977), in that shared relations increased perceived similarity whereas a different relation reduced the perceived similarity of the two analogs.

Second, the results for inductive strength judgments showed a different pattern than that observed for similarity ratings. The G_1G_2P arguments had higher inductive strength than the G_1P arguments (the same pattern as for similarity ratings) but had lower inductive strength than the G_1G_2X and G_1G_2 arguments (the opposite pattern obtained for similarity ratings). These results are consistent with the findings from Experiment 1, confirming that whereas the presence of a shared generative relation increases inductive strength, the absence of a preventive relation increases inductive strength.

In addition, unlike the results for similarity ratings, the G_1G_2X and G_1G_2 arguments were evaluated as having equally high inductive strength. In the similarity ratings, X might have been interpreted as a difference between the source and target, so that the presence of X reduced the similarity. In the inductive strength task, however, X was not related to the effect property, so when a participant was inferring the presence of the effect in the target, X might have been interpreted as a neutral property. Accordingly, the G_1G_2X and G_1G_2 arguments were viewed as equally strong because they involved the same shared causal relations, G_1 and G_2.

The findings from Experiment 2 confirm that people do not simply focus on the number of correspondences between a source and target; rather, they consider the meaning of causal relations and, in particular, whether causes are generative or preventive. By using materials involving causal relations at a higher level of formal complexity, Experiment 2 shows that people use causal models during analogical inference even when the source involves complex higher order relations among multiple objects.

Experiment 3

In the previous experiments, we demonstrated that an analogical inference could be actually strengthened by omitting rather than including in the target a higher order relation (a preventive cause) that is present in the source. One interpretive issue concerns whether our findings directly concern analogical reasoning or might somehow be the indirect result of category-based induction. The close relationship between analogical induction (from individual instances) and category-based induction (from types of instances) suggests caution is warranted. In Experiments 1–2, we manipulated features of animals in a source and target and assumed that participants used analogical reasoning to infer a certain missing property in the target animal on the basis of the properties of the source animal (i.e., analogical inference from one instance to another). However, it is possible that when reasoning about new instances of the well-known category of animals, people actually reasoned at the level of categories, essentially assuming that all individuals of the same animal species share the same basic biological characteristics. Given information about the source animal, people may have naturally projected its properties to the category of the novel species and then in turn applied this categorical knowledge to make inferences about the target animal of the same species (i.e., reasoning from a category to an instance). For example, given a source animal with the structure G_1G_2P, people may infer that these causal relations hold for all animals of that species, in effect setting up a causal model for the species as a whole rather than directly modeling the causal structure of the target individual on the basis of the source.

If our findings are interpreted in terms of category-based induction, then our basic claims still hold at the category level: Categorical inferences appear to be based on causal models and hence are sensitive to the distinction between generative and preventive causes (cf. Ahn, 1999; Lien & Cheng, 2000; Rehder, 2006). Nonetheless, we wished to establish that our findings in fact hold for clear cases of analogical inference, even when category-based induction based on a familiar category is inapplicable.

The materials used in Experiments 1–2 all involved within-domain analogical transfer from one animal to another animal of the same species. In contrast, the core examples of analogical reasoning discussed in the literature involve long-distance, cross-domain transfer between situations based on highly dissimilar entities that nonetheless share relations. Canonical examples include the Rutherford–Bohr model of the atom as an analog of the solar system (Gentner, 1983), the wave theory of sound derived from the behavior of water waves (Holyoak & Thagard, 1995), and the solution to a problem involving radiation therapy based on a military strategy (Gick & Holyoak, 1980). In such cases, it is difficult to claim that any preexisting category is available to allow category-based induction. Rather, the reasoner must directly transfer knowledge from the source in one domain to a target in another. Accordingly, Experiment 3 was designed to examine whether dropping a higher order relation (a preventive cause) present in the source can increase cross-domain analogical transfer.

Method

Participants. Fifty-two undergraduate UCLA students received course credit for participating in the experiment. Participants were randomly assigned to one of two conditions, G_1G_2P or G_1G_2. Twenty-seven participants were assigned to the G_1G_2P condition and 25 participants were assigned to the G_1G_2 condition.

Materials, design, and procedure. Two stories were created, one serving as a source story and the other as a target story, based on two different domains, chemistry and astronomy. A chemist's observations about three liquids and an astronomer's observations about three stars served as the source and target stories, respectively. To ensure that people could not use their prior knowledge of liquids or stars to make inferences, we made all of the liquids and stars novel and imaginary.

Participants first read the source story:

> A research chemist has recently discovered that if three liquid substances, Denitrogel, Oreor, and Tetosium, are mixed together, then a chemical change sometimes occurs so that the molecules of Denitrogel and Tetosium bond together. With molecules of Oreor serving as a catalyst, the molecules of Denitrogel and Tetosium attract each other and the mixed liquid becomes very adhesive, finally changing into a solid material. Through repeated experiments, the scientist has also identified three main factors that determine whether or not the mixed liquids change into a solid.

Following this cover story, three relational observations about the liquids were listed, as summarized in Table 1. Two of the relations tended to cause the effect, changing mixed liquids into a solid, and one of the observations tended to prevent this effect. For example, the fact that "Denitrogel is colder than Oreor" tended to cause formation of a solid, whereas the fact that "The volume of Tetosium is greater than the volume of Denitrogel" tended to prevent formation of a solid. As in the materials used in Experiment 2, the causal relations used in Experiment 3 were formally higher order relations.

After reading the source story, participants read an astronomer's observations about three stars:

> An astronomer who reads about the chemist's findings thinks it may be possible that three stars, Acruxia, Errailel, and Castoriff, located in a distant galaxy, behave in a way similar to the three liquids. The theory is that gravitational attraction among all three stars could make two of the stars move closer together, so that two stars finally fuse to form a super-star. The three stars are close to each other and no other stars have been found in that region of the galaxy.

Following this description of the astronomer's hypothesis, two or three facts about these stars were listed (see Table 1). All of these facts were structurally parallel with the chemist's findings. Table 1 shows the correspondences between the two stories. The corresponding relations across the two stories had varying degrees of semantic overlap. The semantic similarities between relations in the two stories (e.g., "being more turbulent" and "being subject to more violent solar storms") were intended to facilitate the process

Table 1
Schematic Structure of Analogies Used in Experiment 3: Correspondences Between the Chemist's Observations About Three Liquids and the Astronomer's Observations About Three Stars

Structural element	Source	Target
Cover story	Chemist's observations about three liquids	Astronomer's observations about three stars
Generative cause (G_1)	Denitrogel is colder than Oreor	Acruxia has lower temperature than Errailel
Generative cause (G_2)	Oreor is stirred vigorously so it is more turbulent than Tetosium	Errailel is subject to more violent solar storms than is Castoriff
Preventive cause (P_1)	The volume of Tetosium is greater than the volume of Denitrogel	The diameter of Castoriff is wider than the diameter of Acruxia
Effect	Changing into a solid	Forming a super-star

of mapping the two analogs and identifying their relational correspondences.

The experimental design included just two conditions, G_1G_2P and G_1G_2. As in the previous experiments, G and P represent a generative cause and a preventive cause, respectively. In the G_1G_2P condition, two of the facts in the target were semantically and structurally consistent with the corresponding causes of the effect in the source, and one of the target facts was semantically and structurally consistent with the preventive relation in the source. In the G_1G_2 condition, there was no match in the target to the preventive cause in the source. To prevent a blind mapping between the source and target based on the order of listed facts, we randomized the order of the facts for each participant.

After reading the two stories, the chemist's liquid story and the astronomer's star story, two tasks were administered to all participants. We did not collect ratings of similarity in Experiment 3, as it is transparent that the analogs are more similar in the G_1G_2P condition (three shared relations) than in the G_1G_2 condition (two shared relations). We did, however, assess participants' ability to identify the relational correspondences between the two stories. In this mapping task (always administered first), participants were asked to identify which of the three stars (Acruxia, Errailel, or Castoriff) corresponded most closely to each of the three liquids (Denitrogel, Oreor, and Tetosium). This mapping task was intended to ensure that participants read the two stories and to assess whether they could in fact identify structural parallels between the source and the target (a prerequisite for drawing sensible analogical inferences).

The second task involved drawing an analogical inference about the astronomy problem on the basis of the source analog from the domain of chemistry. Participants were asked to judge how likely it was that two of the three stars would fuse to form a super-star. To answer this question, they were told to assume that everything in the descriptions was true and to focus on analogous relations between the chemist's observations and the astronomer's observations. This inductive judgment was made using the same type of frequency scale as was used in the previous experiments. Participants were told to imagine that the astronomer observes 100 cases and to estimate in how many of these cases two of the three stars would fuse to form a super-star, giving a number between 0 and 100.

Results and Discussion

If participants failed to correctly answer all three questions about liquid and star correspondences, we coded their mapping performance as incorrect. Six out of the 52 participants (4 in the G_1G_2P condition and 2 in the G_1G_2 condition) gave incorrect mappings. Because incorrect mappings could have led to erroneous analogical inferences, we first report analyses of inference data excluding these participants. Mean inference judgments for G_1G_2 and G_1G_2P arguments (excluding data from participants who made incorrect mappings) were 69.20 and 52.04, respectively (see Figure 4). These data were analyzed using an independent-samples *t* test, which revealed a significant effect of argument type, $t(44) = 2.59$, $p = .013$. Even when inference data from participants who gave incorrect mappings were included, the effect of argument type remained reliable, $t(50) = 2.20$, $p = .032$.

The results of Experiment 3 thus confirm and extend the basic finding obtained in the previous experiments: When a preventive causal relation present in the source analog is absent from the target, the strength of an inductive inference is correspondingly increased. By demonstrating this effect with cross-domain analogical transfer, the results of Experiment 3 make it clear that this beneficial effect of reducing overlap of a higher order relation is not solely attributable to category-based induction; rather, the phenomenon is observed when people must base an inference about a target analog directly on a single source analog.

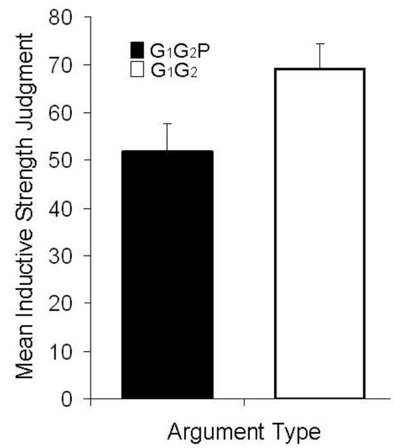

Figure 4. Mean inductive strength judgments for each argument type in Experiment 3. Error bars represent 1 standard error of the mean.

General Discussion

In three experiments reported here, the inductive strength of analogical arguments increased with the number of shared generative causes and decreased with the presence of a shared preventive cause. At the same time, including a preventive cause and the outcome in both analogs necessarily increased the overall match between them (yielding higher rated similarity of the analogs in Experiments 1–2). The positive effect of dropping a shared preventive relation was found both when the causal relations linked properties of a single object (Experiment 1) and when they linked relations between multiple objects (Experiments 2–3). In the latter case, the preventive causal relation was clearly a higher order relation in the formal sense defined by Gentner (1983). Moreover, the effect was found not only for within-domain analogies (Experiments 1–2) but also for cross-domain analogies (Experiment 3). Because cross-domain analogical transfer cannot be accounted for by inferences based on preexisting categories (i.e., category-based induction), our overall findings support the conclusion that analogical inference involves using the source analog to guide construction of a causal model of the target analog. Moreover, given the close linkage between analogical and category-based inferences, this basic conclusion applies to inductive inference broadly, with analogy as an important special case.

In accord with the recent proposal of Bartha (in press), the present experimental findings suggest that people use causal models when they draw analogical inferences. People are likely to first evaluate whether the causal relations in the source are generative or preventive. When mapped to the target, the resulting causal model then provides the basis for inferring the likelihood of a corresponding effect in the target. The presence of a generative cause in the target increases the probability of occurrence of the effect, whereas the presence of a preventive cause decreases the probability of the effect. As a consequence, the presence or absence of a preventive cause in the target has different effects on perceived similarity versus inductive strength: When the target includes the preventive cause, perceived similarity increases but inductive strength decreases. The absence of a preventive cause in the target increases net positive causal power and yields a stronger analogical inference.

The interpretation of analogical inference as being based on causal models is consistent with several findings reported by Lassaline (1996). She found that a shared causal relation that did not involve a given target property did not increase the inductive strength of that property in the target, even though such shared causal relations increased the overall similarity of the analogs. Within a causal model, a generative cause produces its own effects but does not influence the occurrence of causally unrelated properties. Similarly, Lassaline found that the cause relation yielded stronger inductive inferences than did the relation "temporally prior to." The latter relation is a cue for causality (causes typically precede their effects; Lagnado & Sloman, 2004) but by no means guarantees that the relation is causal (e.g., a falling barometer precedes a storm but does not cause it). Hence, "temporally prior to" would be expected to yield weaker inductive inferences than a relation specified to be causal.

Lassaline's (1996) findings are also consistent with the upper bound hypothesis proposed by Bartha (in press). Bartha argued that relations in the source domain influence the evaluation of analogical arguments because the source relations set an upper bound on the causal strength of any analogical inference in the target. That is, cause–effect links inferred in the target are at most as strong as the corresponding cause–effect links in the source. Relations in the source specified as causal will transfer greater causal strengths to links inferred in the target than will noncausal relations such as "temporally prior to."

Implications for Models of Analogical Inference

What is most surprising about the present demonstration that people treat generative and preventive causes differently in evaluating analogical inferences is not the finding itself (described by one reviewer as "so clearly true that it's either deep or trivial") but the fact that it challenges all extant computational models of analogical inference. This fact is yet more surprising because most of the major theoretical discussions of analogy, beginning with Winston (1980), have in one way or another acknowledged the critical importance of causal knowledge as a basis for analogical inference. Gentner (1983) used the cause relation as the central example of a higher order relation, defined as a relation taking another relation as an argument. Her systematicity principle assigns a special status to higher order relations as the basis for constraints on analogical inference. However, both the systematicity principle and its computational implementation, SME (Falkenhainer et al., 1989), base analogical inference solely on the logical form of representations and not on their meaning and hence are committed to the prediction that shared higher order relations can only help and never hinder analogical inferences. The fact that a shared higher order relation—a preventive cause—in fact reduces analogical transfer thus provides a compelling demonstration that any successful model of analogy will need to deal with the meaning of semantic representations, not just their logical form.

Other theorists, such as Winston (1980) and Holyoak (1985), have emphasized the centrality of causal knowledge as a pragmatic (rather than solely structural) constraint on analogical inference. Yet even computational models that allow for the special importance of causal relations are in their current implementations unable to account for our present findings. For example, both ACME (Holyoak & Thagard, 1989) and LISA (Hummel & Holyoak, 2003) include mechanisms for placing greater weight on causal relations so that these have especially strong influences on analogical mapping and inference. In their current implementations, ACME and LISA would naturally mark both generative and preventive causes as important and focus attention on them. However, these models are unable to grasp that the two types of causes are important in different ways, with very different implications for analogical inference—a preventive cause is indeed important, but the outcome would be more probable without it. Both ACME and LISA compute a match score, which, although based on more than structural overlap alone (unlike SME), nonetheless is constrained to increase monotonically with the number of matching relations. The match score in turn determines the probability of an analogical inference; hence none of these models are able to account for the decrease in inductive strength that results from a match between preventive causes present in both the source and the target analogs.

The basic problem for all extant models of analogy is that they lack a detailed representation of causal knowledge that could support commonsense causal reasoning (e.g., generative causes make things happen, but preventive causes stop things from happening; causes produce effects, but effects do not produce their causes; multiple causes combine according to specific generating functions; see Buehner & Cheng, 2005; Waldmann, 2007; Waldmann & Holyoak, 1992). Thus the commonsense nature of the present findings highlights the weakness of commonsense reasoning in current models of analogy.

Analogical Reasoning With Interpreted Structures

The present findings add to a growing body of work suggesting that future theoretical progress in understanding analogical reasoning requires taking a broader view of the overall process. As pointed out by Bassok, Wu, and Olseth (1995), analogy models have typically focused on the mechanisms by which correspondences can be established between predetermined representations of a source and target and then exploited to generate candidate inferences about the target. Both the input representations (source and target) and the output (target inferences) are treated as static. But as Bassok et al. argued, human analogical reasoning appears to be based on highly dynamic representations. The critical role of semantic interpretation in analogical processing is bolstered by evidence that analogical processing automatically activates relevant category relations (Green, Fugelsang, & Dunbar, 2006; Green, Fugelsang, Kraemer, & Dunbar, 2008). The initial representations of the source and target may undergo semantic interpretation and elaboration that in turn affect the ease of analogical transfer. For example, Bassok et al. observed that people expect objects to be assigned to people (e.g., golf carts assigned to caddies) rather than the reverse. These investigators demonstrated that when a source analog includes an ambiguous or incongruous assign relation (e.g., caddies assigned to carts), the stated relation is likely to be reinterpreted to bring it into accord with prior expectations. The ease or difficulty of subsequent mapping and transfer then depends on the correspondences between the interpreted representations (see also Kotovsky, Hayes, & Simon, 1985). Although some efforts have been made to model the dynamic aspects of analog interpretation by treating mapping as an iterative process (Hofstadter & Mitchell, 1994; Hummel & Holyoak, 2003), such work has been limited in scope.

Whereas research such as that described above has focused attention on the dynamic nature of the input to analogical processing, the present findings call attention to the dynamic nature of its output. Extant analogy models apply a CWSG algorithm to a computed mapping to produce a set of inferences about the target. These inferences, like the analogs themselves, are represented in some static code. The degree of belief in an inference is solely determined by some measure of the goodness of the source–target mapping from which the inference was derived. What is lacking, given the findings reported here, is some mechanism for dynamically updating the degree of belief in an analogical inference on the basis of the causal relationships within the target itself.

Moreover, our findings suggest that the missing theoretical mechanism for dynamic inference evaluation cannot be simply outsourced to some postanalogical module, such as verification based on direct knowledge acquired about the target (as discussed, e.g., by Falkenhainer et al., 1989, p. 23). Consider the situation confronting the reasoner in our Experiment 3: No direct knowledge about the behavior of the imaginary star system that forms the target domain is ever provided beyond that knowledge allowing computation of the analogical inference in the first place. That is, the reasoner has no additional basis for assessing how likely it is that two stars will fuse to form a super-star. Rather, the causal model of the target—constructed by analogical inference from the source domain of chemical reactions—itself allows dynamic evaluation of degree of belief in the candidate inference. As Bartha (in press) argued on the basis of actual examples from the history of science, the credibility of analogical inferences can be assessed in part by internal criteria. In nonexperimental sciences, such as astronomy and archaeology, analogy may, at times, provide the most direct source of knowledge available for evaluating inferences about the target.

Integrating Causal Models With Analogical Inference

The present findings suggest that future theoretical work should aim to build models of analogy that incorporate the basic elements of causal models (e.g., Cheng, 1997; Griffiths & Tenenbaum, 2005; Pearl, 1988; Waldmann & Holyoak, 1992). This remains a challenging enterprise. One impediment is that there is a representational gap between what are termed *causal Bayes nets* (Pearl & Russell, 2003), as exemplified by Figure 1, and the predicate-calculus style representations that underlie most major analogy models. Although we have informally characterized the type of causal arrows shown in Figure 1 as representing higher order relations, causal Bayes nets typically are interpreted as lacking the representational power of predicate calculus. Thus a factor such as G_1 is simply an unanalyzed node in a Bayes net, rather than a full-blown propositional structure such as "Denitrogel is colder than Oreor" (Experiment 3). From the point of view of models of analogy, representing the internal structure of propositions is essential to the fundamental operation of analogical mapping and inference. Thus a basic requirement for integrating causal Bayes nets with models of analogy is to extend the representational power of the former (for efforts in this direction, see Milch et al., 2007).

The potential for fruitful interplay between models of analogy and of causal inference goes in both directions. In particular, analogical inference provides a possible answer to a central question that looms in current work on causal models, namely, how are causal hypotheses first formed? One suggestion has been that causal hypotheses are generated by some form of causal grammar (Griffiths & Tenenbaum, 2007). When a source analog is available, an alternative possibility is that the CWSG procedure for building relational structures by analogy can serve to build an initial causal model of the target. An effective analogy operates by quickly identifying a high probability region in what may be a vast search space, generating causal hypotheses about the target that are more likely (relative to blind search) to be approximately correct.

A study of problem solving by Schunn and Dunbar (1996) illustrates how cross-domain analogical transfer may guide formation of a causal model. In an initial session involving a problem in biochemistry, some participants learned that addition of an inhibitory enzyme (preventive cause) decreased virus reproduction. In a subsequent session the following day, these same participants were asked to solve a molecular-genetics problem, which also involved a preventive cause (an inhibitory gene). Schunn and Dunbar found that participants who had been exposed to the concept of inhibition in the initial session were more likely than control participants to develop a solution based on inhibition for the transfer problem, even though experimental participants evinced no signs of awareness that the earlier virus problem had influenced their solution to the gene problem. These findings suggest that some form of analogical transfer (perhaps implicit) can guide the construction of a causal model appropriate for the target domain. More generally, analogy and causal inference are intricately related (Holland et al., 1986), and a full theory of human induction will need to provide a unified account of both.

References

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

Bartha, P. (in press). *By parallel reasoning: The construction and evaluation of analogical arguments.* Oxford, United Kingdom: Oxford University Press.

Bassok, M., Wu, L.-L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition, 23,* 354–367.

Buehner, M., & Cheng, P. W. (2005). Causal learning. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 143–168). Cambridge, United Kingdom: Cambridge University Press.

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1119–1140.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

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

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

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

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

Green, A. E., Fugelsang, J. A., & Dunbar, K. N. (2006). Automatic activation of categorical and abstract analogical relations in analogical reasoning. *Memory & Cognition, 34,* 1414–1421.

Green, A. E., Fugelsang, J. A., Kraemer, D. J., & Dunbar, K. N. (2008). The micro-category account of analogy. *Cognition, 106,* 1004–1016.

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

Griffiths, T. L., & Tenenbaum, J. B. (2007). Two proposals for causal grammars. In A. Gopnik & L. Schultz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 323–345). Oxford, United Kingdom: Oxford University Press.

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

Hofstadter, D. R., & Mitchell, M. (1994). The copycat project: A model of mental fluidity and analogy-making. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 31–112). Norwood, NJ: Ablex.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

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., 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* (pp. 113–180). Norwood, NJ: Ablex.

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

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

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

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110,* 220–264.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17,* 248–294.

Lagnado, D. A., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 856–876.

Lassaline, M. E. (1996). Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 754–770.

Lee, H. S., & Holyoak, K. J. (2007). Causal models guide analogical inference. In D. S. McNamara & G. Trafton (Eds.), *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society* (pp. 1205–1210). Austin, TX: Cognitive Science Society.

Lien, Y., & Cheng, P. W. (2000). Distinguishing genuine from spurious causes: A coherence hypothesis. *Cognitive Psychology, 40,* 87–137.

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2006). Modeling causal learning using Bayesian generic priors on generative and preventive powers. In R. Sun & N. Miyake (Eds.), *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society* (pp. 519–524). Mahwah, NJ: Erlbaum.

Lu, H., Yuille, A., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (in press). Bayesian generic priors for causal learning. *Psychological Review.*

Milch, B., Marthi, B., Russell, S., Sontag, D., Ong, D. L., & Kolobov, A. (2007). BLOG: Probabilistic models with unknown objects. In L. Getoor & B. Taskar (Eds.), *Introduction to statistical relational learning* (pp. 373–398). Cambridge, MA: MIT Press.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185–200.

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

Pearl, J., & Russell, S. (2003). Bayesian networks. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 157–160). Cambridge, MA: MIT Press.

Read, S. J. (1984). Analogical reasoning in social judgment: The importance of causal theories. *Journal of Personality and Social Psychology, 46,* 14–25.

Rehder, B. (2006). When similarity and causality compete in category-based property generalization. *Memory & Cognition, 34,* 3–16.

Rehder, B. (2007). Property generalization as causal reasoning. In A. Feeney & E. Heith (Eds.), *Inductive reasoning: Experimental, developmental, and computational approaches* (pp. 81–113). Cambridge, UK: Cambridge University Press.

Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology, 50,* 264–314.

Reichenbach, H. (1956). *The direction of time.* Berkeley: University of California Press.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14,* 665–681.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–59). New York: Cambridge University Press.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world.* Princeton, NJ: Princeton University Press.

Schunn, C. D., & Dunbar, K. (1996). Priming, analogy, and awareness in complex reasoning. *Memory & Cognition, 24,* 271–284.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25,* 231–280.

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

Sloman, S. A. (1997). Explanatory coherence and the induction of properties. *Thinking & Reasoning, 3,* 81–110.

Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition, 22,* 377–386.

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

Tenenbaum, J. B., Kemp, C., & Shafto, P. (2007). Theory-based Bayesian models of inductive reasoning. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Experimental, developmental, and computational approaches* (pp. 167–204). Cambridge, United Kingdom: Cambridge University Press.

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

Waldmann, M. R. (2007). Combining versus analyzing multiple causes: How domain assumptions and task context affect integration rules. *Cognitive Science, 31,* 233–256.

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.

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

Received October 25, 2007
Revision received April 15, 2008
Accepted April 22, 2008 ■