Causal Asymmetry in Inductive Judgments

Philip M. Fernbach (philip_fernbach@brown.edu)
Adam Darlow (adam_darlow@brown.edu)
Brown University, Department of Cognitive and Linguistic Sciences, Box 1978
Providence, RI 02912 USA

Abstract

We propose a normative model of inductive reasoning about *causal arguments*, those in which there is a direct causal relation between categories. The model derives inductive judgments from a causal Bayesian network that represents the causal structure of the argument. It supports inferences in the *causal direction* (e.g. a mother is drug-addicted, how likely is it that her newborn baby is drug-addicted?), and in the *diagnostic direction* (e.g. a newborn baby is drug-addicted, how likely is it that the baby's mother is drug-addicted?). We explored how causal and diagnostic judgments should change as a function of the parameters of the model, which include the prior probability of the cause, the causal power of the cause to bring about the effect, and the strength of alternative causes. The model was fit to the results of an experiment in which we manipulated the strength of alternative causes by varying the predicate while keeping the categories constant. Contrary to the predictions of previous theories, participants were not biased to over-estimate causal judgments relative to diagnostic judgments. Instead, they neglected alternative causes when reasoning causally and hence systematically *underestimated* causal judgments. Conversely, diagnostic judgments were sensitive to the strength of alternative causes and were unbiased, demonstrating that inductive reasoning is sensitive to some rational principles.

Key Words: Inductive Inference, Causal Reasoning, Diagnostic Reasoning, Causal Models, Causal Bayesian Networks, Category-Based Induction, Probabilistic Models

Introduction

Causal knowledge is central to many inductive inferences (For a review see Sloman and Lagnado, 2005). The most direct illustration comes from studies of causal arguments, those in which there is a relation of transmission between categories, for example the transmission of a drug-addiction between a mother and her newborn. Inductive judgments that require reasoning from cause to effect (e.g. the probability that a newborn has a drug addiction given that its mother does) are referred to as *causal judgments* while judgments that require reasoning from effect to cause (e.g. the probability that a mother has a drug addiction given that her newborn does) are referred to as *diagnostic judgments.*

Previous theories of inductive reasoning about causal arguments propose that it is natural for people to reason from causes to effects but not from effects to causes. Hence causal judgments are overestimated relative to diagnostic ones, *ceteris paribus*. A prominent example comes from Kahneman and Tversky (1982) who report that participants rated the likelihood that a daughter has blue eyes given that her mother does to be higher than the likelihood that a mother has blue eyes given that her daughter does. They argue that the normative probabilities are equal because the base rate probability of blue eyes should be equal across generations and therefore the conditional probabilities should also be equal. Medin et al. (2003) propose a similar idea. Their relevance framework predicts a *causal asymmetry* in judgments because it is easier to reason from causes to effects than from effects to causes. For instance, the likelihood of lions having a property given that hyenas have it is higher than the likelihood of hyenas having it given that lions do because there is a relation of transmission from hyenas to lions through the food chain. Unlike Kahneman and Tversky, they do not analyze the normative force of their claim.

The purpose of this paper is twofold. First we describe a simple probabilistic model of causal arguments that allowed us to derive inductive judgments from an abstract representation of the causal structure of an argument. The model is based on causal Bayesian networks, representations of causal structure that obey the laws of probability (Spirtes, Glymour & Scheines, 1993; Pearl, 2000). The rational basis of the model supports a normative analysis of inductive judgments. Second, we report the results of an experiment that assessed human inductive reasoning in light of the normative analysis. Unlike previous experiments, we collected conditional probability judgments (i.e. the causal and diagnostic judgments) along with the primitives for those conditionals (i.e. the model parameters). This allowed us to fit the model based on the parameters we collected and compare those fits to the causal and diagnostic judgments that were probed directly.

According to the analysis, one important determinant of the relative strength of causal and diagnostic strength is the strength of alternative causes. This prediction was tested in the experiment. The results showed that diagnostic judgments were sensitive to the strength of alternative causes and unbiased, while causal judgments were systematically underestimated due to the neglect of alternatives.

Causal Model of Inductive Judgments

Model Description Causal Bayes nets are graphs with nodes that represent properties or events and edges that represent the causal relations among the nodes. They can be used to compute the probabilities of unobserved nodes given observation of or interventions on other nodes (Pearl, 2000). Causal Bayes nets are suited to modeling normative causal judgments because they combine probabilistic inference with an interventional logic, and intervention is the hallmark of causality (Woodward, 2004). A transmission argument can be represented by a common-effect structure, one in which there are multiple possible causes for an effect. In general, a predicate might be transmitted to the effect category from the target cause, or by some alternative cause. To capture the additional constraint that a true alternative cause should be independent of the target cause we restrict ourselves to arguments in which transmission from a source to a recipient follows an independent causal path and use a noisy-or function to specify how causes combine. The presence of either cause raises the probability of the effect and if both causes are present the probability of the effect is even higher, increasing according to the independent contribution of each cause.

Model Description A causal Bayes net can be fully described by the probability distributions of the exogenous variables, those that have no parents in the graph, along with a set of functions and parameters that define the probability distributions of endogenous nodes conditioned on their parents. In other words, the model requires specifying the prior probability distributions of the cause and the alternatives and a function describing how the cause and the alternatives combine to generate the effect.

By aggregating all alternative causes into a single node, a causal background (Cheng, 1997), the structure necessary for defining causal and diagnostic probabilities can be concisely represented as a causal Bayes net with three nodes: the cause, the effect and the aggregate of all alternative causes. Separate edges connect the cause and alternative to the effect. To specify the parameters over this structure we assumed that events are binary; they either happen or they do not. This allowed us to represent the probability distribution of exogenous nodes with a single number, a prior probability. We also assumed that the cause and any alternative causes are independent and generate the effect independently according to a nosiy-or function as discussed above. The independent contribution of a cause can be defined in the model as a parameter that specifies the conditional probability of the effect given that cause and no others (a 'causal power'). Because of its use of the noisy-or function and parameterization in terms of causal powers, the structure is identical to that proposed in Cheng's seminal PowerPC model of causal learning.

To simplify calculations, we collapsed the prior probability and causal power of the alternative causes into a single parameter denoting the strength of alternatives, set to *P(Effect | ~Cause)*. This is akin to setting the prior to one (i.e. assuming alternatives are always present but only effective in bringing about the effect some of the time.) The prior and causal power of alternatives are always confounded in the model, so the simplification is not substantive.

The model is therefore fully parameterized by three numbers: the prior probability of the cause (P_c), the causal power of the cause (W_c) equal to *P(Effect | Cause, ~Alternative Causes),* and the strength of alternatives (W_a) or *P(Effect | ~Cause)*. The structure and parameterization are depicted in Figure 1. In the figure W_a represents both the prior and causal power of alternatives collapsed into a single term.

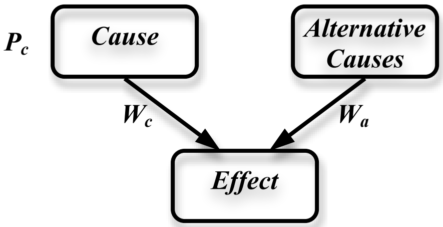

Figure 1: Bayes net model of causal arguments

The causal judgment (C) and diagnostic judgment (D) correspond to *P(Effect|Cause)* and *P(Cause|Effect)*, respectively. C is calculated by direct application of the noisy-or equation:

$$C = P(Effect \mid Cause) = W_c + W_a - W_c W_a \quad (1)$$

Note the difference between W_c and C. C represents the probability that the effect occurs given that the cause occurred. This includes the cases in which the cause was effective in generating the effect, but it also includes cases in which the cause was ineffective but an alternative cause was effective. Therefore, C is higher than W_c and it increases with the strength of alternatives.

The diagnostic judgment, D, is derived by considering it's complement, the probability that the cause *did not* occur despite the effect having occurred.

$$D = P(Cause \mid Effect) = 1 - P(\sim Cause \mid Effect) \quad (2)$$

By Bayes' rule:

$$D = 1 - P\left(Effect \mid \sim Cause\right)\frac{P\left(\sim Cause\right)}{P\left(Effect\right)} \quad (3)$$

Deriving *P(Effect)* by the noisy-or equation and substituting W_a for *P(Effect|~Cause)* and *(1 - P_c)* for *P(~Cause):*

$$D = 1 - (1 - P_c)\frac{W_a}{P_c W_c + W_a - P_c W_c W_a} \quad (4)$$

Equation 4 shows that two factors determine D: the prior probability of the cause and the probability that the alternatives caused the effect (i.e. the ratio between W_a and the expansion of *P(Effect)* at the end of Equation 4). The presence of the effect cannot decrease the probability of the cause, so D is always higher than P_c and it increases with P_c. Conversely, the effect is diagnostic of the cause to the extent it was not generated by alternative causes. Therefore, the cause and the alternatives compete to explain the effect and D decreases with the probability that the alternative causes caused the effect.

Model Predictions C is a function of two parameters, W_c and W_a, and increases as each of them increases independently. D is a more complex function of all three parameters. As mentioned earlier, it depends on the prior probability of the cause and the probability that the effect was caused by the alternatives. The probability that the effect was caused by the alternatives is a comparative measure of the strength of alternatives relative to the strength of the cause. Accordingly, it increases with W_a and decreases with P_c and W_c. Therefore, D increases as P_c or W_c increases or as W_a decreases.

Experiment

The experiment was motivated by the normative analysis. First we wanted to vary the strength of alternative causes to see whether people are sensitive to that factor as prescribed by the model. For each set of categories we generated two predicates, one that suggested strong alternatives and one that suggested weak or absent alternatives. For instance, there are no strong alternative causes to a baby's drug-addiction besides the mother, but dark skin could be transmitted by the father. We predicted that diagnostic judgments would be stronger for weak alternative items despite the inference being about the same categories.

Second, we wanted to generate enough data to fit the model. We therefore collected causal and diagnostic judgments and the model parameters P_c, W_c and W_a. To probe these we simply asked for the likelihood of the relevant events on a 0-100 scale. Examples of the question forms are shown in Table 1. If people's inductive judgments are consistent with their beliefs about the relevant probabilities then the conditional probabilities derived from the parameters should match the causal and diagnostic judgments.

Our basic method was to rely on pre-existing beliefs about the causal structures and probabilities rather than train people on novel causal systems (e.g. Rehder, 2006). The benefit of this approach was that the experimental method was streamlined, necessitated no training and was intuitive for participants. One possible concern was that differences across conditions could be driven by beliefs about particular predicates or categories or by items that did not perfectly reflect the modeling assumptions. We therefore used a large number of arguments, a weak and strong version for each of 20 sets of categories, 40 in all.

Table 1: Example Question Forms

Parameter / Judgment	*Example Wording*
Prior Probability of Cause (P_c)	A woman is the mother of a newborn baby. How likely is it that the woman is drug-addicted?
Causal Power of Cause (W_c)	The mother of a newborn baby is drug-addicted. How likely is it that her being drug-addicted causes her baby to be drug addicted?
Strength of Alternatives (W_a)	The mother of a newborn baby is not drug addicted. How likely is it that her baby is drug addicted?
Causal Judgment (C)	The mother of a newborn baby is drug-addicted. How likely is it that her baby is drug-addicted?
Diagnostic Judgment (D)	A newborn baby is drug addicted. How likely is it that its mother is drug addicted?

Method

Participants 162 participants were recruited by Internet advertisement and participated online for the chance to win a $100 lottery prize. Additionally, 18 Brown University students participated in the lab for class credit or were paid at a rate of eight dollars per hour. In total 180 Participants completed the experiment. Each participant was randomly assigned to one of five groups.

Design The experiment had 3 independent variables: categories, strong versus weak alternatives and question type. Each set of categories consisted of a cause and an effect category where the predicate could be transmitted to the effect by the cause. For each set of categories we generated two predicates, one that implied strong alternative causes for the possession of the predicate by the effect category and one that implied weak alternative causes. Categories and predicates were chosen to fit the common effect noisy-or causal structure where any alternative causes provide an independent contribution to the effect and the causal relation from cause to effect is unidirectional. For each predicate we asked five questions, the prior probability of the cause (P_c), the causal power of the cause (W_c) the strength of alternatives (W_a), the causal judgment (C) and the diagnostic judgment (D). We chose 20 sets of categories, two predicates for each set, and five questions for each predicate for a total of 200 questions.

To avoid interactions among questions about the same predicate, the variables were manipulated in the following way: We split the 200 questions into five questionnaires with 40 questions each. Questions were randomly assigned such that each questionnaire had one question type from each of the 40 predicates and so that no questionnaire had the same question type for the weak and strong version of a given set of categories. Each participant therefore answered a single question about each predicate. The order of questions in each questionnaire was randomized and was identical across participants assigned to that questionnaire.

Materials and Procedure Examples of some of the categories and predicates used in the experiment are shown in Table 2. The experiment was completed on a computer either in our lab or offsite. Participants were randomly assigned to one of the five questionnaires. Each questionnaire consisted of instructions at the top followed by 40 questions, all on a single sheet. The instructions read, "please estimate how likely the following events are from the small amount of information given to you. Give an answer between 0 (impossible) and 100 (definite) in the space provided for each of the questions. Don't think too hard about each one as there is no correct answer but don't guess wildly either." After completing the questions, participants clicked a button to submit their form. The experiment took approximately 20 minutes.

Table 2: Examples of Predicates and Categories Tested in the Experiment

Cause Category	*Effect Category*	*Strong Alternatives Predicate*	*Weak Alternatives Predicate*
Mother	Newborn baby	Has dark skin	Is drug-addicted
Coach	High school football team	Is motivated	Knows a complicated play
Mayor of a major city	New Policy	Is unpopular	Is fiscally conservative
Apple Slices used to make an apple pie	Apple Pie	Are sweet	Have seeds
Music at a party	Party	Is loud	Is good for dancing
Transfusion blood at African Hospital	Transfusion Patient	Has an infectious disease	Is anemic
Engine of a 2005 Honda accord	2005 Honda Accord	Is not functioning properly	Smells of burnt oil
Body of water	Stew made from fish from the body of water	Is salty	Is high in mercury

Results

Five participants gave the same response to each question and were omitted from subsequent analysis. The mean causal and diagnostic judgments for the strong and weak alternatives conditions are shown in Figure 2. We collapsed the data across participants and assessed the relative effect of strength of alternatives on causal and diagnostic judgments by performing a 2 (alternatives: strong vs. weak) x 2 (judgment: causal vs. diagnostic) repeated measures ANOVA. There was a significant interaction between judgment type and strength of alternatives ($F(1,19)=30.71$, $p=0$). There was also a main effect of strength of alternatives ($F(1,19)=4.96$, $p=0.038$) but no significant effect of type of judgment ($F(1,19)=0.65, p=0.43$).

We conducted planned comparisons between judgments in the strong and weak alternatives conditions, which revealed that diagnostic judgments in the weak alternatives condition (M = 81.7) were higher than in the strong alternatives condition (M= 58.5; $t(19) = 4.95$, $p=0$). As in Experiment 1, causal judgments did not differ significantly ($M_{strong} = 75.3$; $M_{weak} = 69.6$; $t(19) = 1.31$, $p=0.24$). We also used matched sample t-tests to compare mean parameter judgments for each category set across the strong/weak manipulation. The results are shown in Table 3. The manipulation of strong vs. weak alternatives was effective as evidenced by the difference between W_a in the two conditions. P_c and W_c responses didn't differ significantly between conditions.

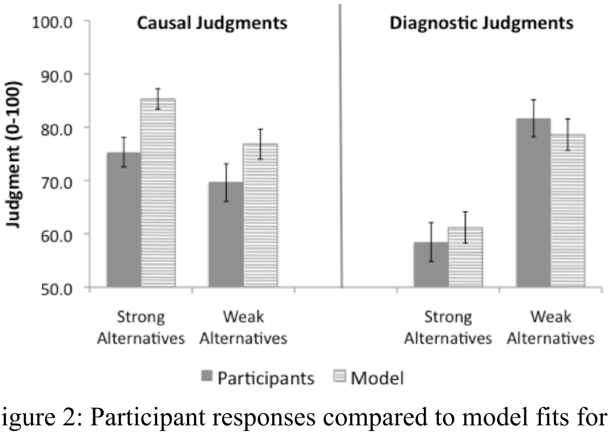

Figure 2: Participant responses compared to model fits for the strong/weak alternatives conditions. Causal Judgments are on the left and diagnostic judgments on the right.

Model Fits

The model represents the relation between a single participant's judgments of the parameters P_c, W_c and W_a and their judgments of C and D. Because of the incomplete design, no participant made all of the parameter judgments for any single item, and we therefore had a distribution of unmatched judgments of the parameters for each item. We could not simply take the means of these distributions and combine them according to the model's Equations because it is not generally true that the mean of a function of distributions is equivalent to applying that function to their means. In particular, the equation for D, which includes terms in the denominator, violates this assumption. For C the assumption did hold, and the model's outputs for C were the same as if they were calculated directly from the parameter means. Nonetheless, for consistency's sake we used the same procedure to generate predictions for C and D.

Our method was to use a sampling procedure to generate a distribution for the model's predictions of C and D for each item and used the mean of this distribution as the model's prediction for that item. To generate a single sample of C and D for a given item we drew one sample of each of the three parameters uniformly and independently from the set of participant responses. We then calculated C and D from the sampled parameters according to Equations 1 and 4. We repeated this procedure to generate 100,000 samples each of C and D for each item and took the means as the model's predictions for that item. Reruns of the sampling procedure yielded no differences in the predictions for either C or D.

Modeling Results Figure 2 shows the model predictions for C (left panel) and D compared to participant responses. As with participant responses, model predictions for D were higher in the weak condition (M=0.79) than in the strong condition (M=0.61; $t(19)= 4.98$, $p=0$). Model predictions for C were lower in the weak condition (M=0.77) than in the strong condition (M=0.85; $t(19)=2.38$, $p=0.028$). The model predictions of D were not significantly different from participant responses ($t(39)=0.67$, $p=0.94$) and were highly correlated with items in the strong and weak conditions separately ($r_{strong}=0.69$, $p=0$; $r_{weak}=0.69$, $p=0$) and across both conditions ($r=0.80$, $p=0$). Model predictions of C (M=0.81) were significantly higher than participant responses (M=0.72; t(39)=6.54, p=0), but were still highly correlated both within each condition ($r_{weak}=0.83$, $p=0$; $r_{strong}=0.75$, $p=0$) and across conditions ($r=0.72$, $p=0$).

To test whether the full model was necessary to predict participants' responses we ran multiple regression analyses on the fits to C and D. For judgments of D we considered the possibility that the high correlation between the model and judgments of D could be driven primarily by differences in W_a. W_a was significantly correlated with D across the strong/weak manipulation (r=-0.49, p=0.0007), however the correlations were not significant in each condition separately (r_{weak}=-0.28, p=0.23; r_{strong}=-0.08, p=0.74). The multiple regression, which used W_a and the full model as predictors, showed that the model fit the data better than W_a alone and W_a had no predictive value beyond its role in the model. Together, the full model and W_a accounted for 64% of the variance in D. The unique variance of the full model accounted for 41% of the variance of D (t=6.46, p=0), but the unique variance of W_{a} did not account for any of the variance of D (t=1.00, p=0.32).

In contrast, the best predictor of causal judgments was the single parameter W_c and not the full model. W_c alone fit the data better than the model and the model had no predictive value beyond that of W_c. The model and W_c together accounted for 77% of the variance of C. The unique variance of Wc accounted for 10% of the variance of C (t=4.14, p=0), but the unique variance of the model did not account for any of the variance of C (t=0.64, p=0.53). Because W_c and W_a are the only two factors in the model prediction of C these results imply that causal judgments were independent of W_a, which we verified (r=0.044, p=0.78). Corroborating this analysis we also found that there was no significant difference between judgments of C and W_c (t(39)=0.60, p=0.55).

Discussion

The implications of the experiment have a dual nature. On one hand, the normative analysis identified a rational principle that is not accommodated by other theories, namely the effect of alternative strength on inductive judgments. The strong/weak manipulation had a large effect on the relative strength of causal and diagnostic judgments as prescribed by the normative analysis, showing that inductive reasoning is sensitive to some of the factors that it should be. Similarity-based models (Osherson et al., 1990; Sloman, 1993) make no prediction regarding predicate differences while the relevance framework (Medin et al, 2003) predicts asymmetry favoring the causal direction only. When assessed in light of our model the asymmetry reported by Medin et al. may be rational if the predicates they used implied strong alternatives.

On the other hand, the model helped us identify a violation of normative judgment. Participants systematically neglected alternative causes when reasoning causally but not when reasoning diagnostically. Causal judgments were lower than they should have been if reasoning were perfectly consistent. Diagnostic judgments were unbiased.

While our analyses do not support previous theories that claim a bias in judgments of probability in the causal direction, they do share some common assumptions. For instance we chose to parameterize our model using causal powers, which have an inherent causal directionality. The result of this choice was that the diagnostic judgment derived from the model is more complex than the causal judgment in the sense that it contains more terms. Kahneman and Tversky (1982) reported that participants were more confident in causal judgments than diagnostic ones. One explanation of this finding is that people's inductive thinking draws on knowledge about how causes generate effects even when thinking diagnostically. Thus causal judgments are relatively simple functions of knowledge that is readily available, like causal powers, while diagnostic judgments are more indirect (also see Krynski & Tenenbaum, 2007). Contrary to Kahneman and Tversky, our results indicate that the outcome of this preference is not a bias to overestimate causal judgments, but instead a bias to underestimate them.

Follow-Up Work

Follow up work has addressed several possible alternative explanations for our results. Space does not permit a full account of this work but two particularly important issues bear mentioning. First, one might ask whether the results are contingent on the fact that the data were collected with a design that was primarily between subjects and that model fits for diagnostic judgments were obtained by simulation. We reran a modified version of the experiment fully within subjects and the results corroborated the phenomena reported here. Second, the conclusion that participants neglected alternatives in the causal direction is based primarily on the high correlation between causal judgments and judgments of causal power, W_c. One possible alternative interpretation of this finding is that this correlation reflects how people interpret the probe questions and not a phenomenon of reasoning. In other words, participants may simply be interpreting the causal question as asking for a judgment of causal power or conversely, interpreting the W_c question as asking for the causal judgment. We tested this possibility by mentioning an alternative cause explicitly and then asking the W_c and causal questions. If participants understand the difference between the two but neglect alternatives then judgments should be the same when alternatives are not mentioned. When they are mentioned, judgments of W_c should stay the same while causal judgments should increase. This is precisely what we found and provides evidence against the linguistic interpretation.

The neglect finding is further supported by a more direct manipulation. Romeo et al. (2008) asked medical professionals to reason causally and diagnostically about the presence of diseases and symptoms. When asked to judge the probability of a symptom given a disease, participants gave the same response as when asked to judge the probability of a symptom given the disease and a test that showed that no other diseases were present. However when asked the corresponding diagnostic questions (the probability of disease given symptom versus the probability of disease given symptom and no other disease) the responses were very different indicating that the medical professionals neglected alternatives when reasoning causally but not diagnostically.

The finding of neglect is perhaps not surprising in light of similar findings that people tend to focus on few or singular explanations such as when they are asked to generate explanations for why an automobile failed to function properly (Mehle, 1982). The novel perspective on these findings that our work provides is that neglect disappears when people are asked for a diagnostic judgment as opposed to generating explanations or making predictive judgments.

Our speculative hypothesis is that this is a result of inductive reasoning's reliance on heuristics that aim to minimize effort. Because of the structure of common cause networks of the type used in our model, the causal judgment can be reasonably approximated using the causal power of the main cause, assuming it is relatively strong, while ignoring alternatives. This is not true of diagnostic judgments which are inherently comparative in the sense that they measure how likely the target cause was to have brought about the effect relative to other causes. To completely neglect alternatives when reasoning diagnostically would be to give a diagnostic strength of one, or to assume that the target cause must be present given that the effect is. This would not be a very useful strategy for diagnostic reasoning.

Related Model

Shafto et al. (in press) propose an alternative approach to modeling inductive judgments about causal arguments. In their model, inductive strength is a function of a Bayesian inference that computes the posterior probability of different groupings of categories conditioned on the evidence and the prior distribution over category groupings is derived from a structured representation of the categories. Shafto et al.'s hypothesis is that the structure that is used for an inference is a function of the predicate. For example, disease predicates bring to mind relational structures that support transmission, like food chains, while genetic predicates bring to mind hierarchical taxonomies. The model predicts the same asymmetry as Medin et al. (2003) when reasoning about diseases but not genes. Our model is not strictly comparable because we use a local structure that applies to a broad range of causal arguments regardless of their global structure, and the parameters for a given predicate are inputs into the model. In Shafto et al.'s model the parameters are derived from more abstract information, the domain-specific relational structure associated with the predicate.

Our empirical findings provide at least one challenge to their approach. The neglect finding shows that people don't always take into account all the information that they should when reasoning inductively (also see Sloman, 1998). Their idea is inconsistent with this because inference is mediated by relational structures that represent all the categories in a domain even when reasoning about just a subset of those categories.

Acknowledgments

This work was supported by NSF Award 0518147. We thank Steven Sloman for guidance on the project and Jonathan Bogard for help collecting data.

References

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

Krynski, T. R. & Tenenbaum, J. B. (2007) The role of causality in judgment under uncertainty. Journal of Experimental Psychology: General, 136 (3), 430-450.

Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. K. (2003). A relevance theory of induction. Psychonomic Bulletin and Review, 10 (3), 517-532.

Mehle, T. (1982) Hypothesis generation in an automobile malfunction inference task. Acta Psychologica, 52, 87-106.

Osherson, D. M., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. Psychological Review, 97, 185-200.

Pearl, J. (2000). Causality. Cambridge: Cambridge University Press.

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

Romeo, S., Sutton-Skinner, K., Petersen, T., Baer, L., Huffman, J., Birnbaum, R. & Sloman, S. A. (2008). Clinical decision making biases in a group of mental health providers. Poster presented at the Simches Symposium, Boston, MA.

Shafto, P., Kemp, C., Baraff Bonawitz, E., Coley, J. D. & Tenenbaum, J. B. (in press). Inductive reasoning about causally transmitted properties. Cognition.

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

Sloman, S. A. (1998). Categorical inference is not a tree: The myth of inheritance hierarchies. Cognitive Psychology, 35, 1-33.

Sloman, S. A., & Lagnado, D. A. (2005). The problem of induction. In R. Morrison & K. Holyoak (Eds.), Cambridge handbook of thinking and reasoning (95-116). New York: Cambridge University Press.

Spirtes, P., Glymour, C. & Scheines R. (1993). Causation, prediction and search. New York: Springer-Verlag.

Tversky, A. & Kahneman, D. (1982). Causal schemas in judgements under uncertainty. In D. Kahneman, P. Slovic & A. Tversky (eds.), Judgement under uncertainty: Heuristics and biases (117-128). Cambridge: Cambridge University Press.

Woodward, J. (2003). Making things happen: A theory of causal explanation. New York: Oxford University Press.