Exclusivity in causal reasoning

Alexander LaTourrette (latourrette@u.northwestern.edu)
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

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

Lance J. Rips (rips@northwestern.edu)
Department of Psychology, 2029 Sheridan Road
Evanston, IL 60208 USA

Abstract

Causal systems often include mutually exclusive events: events which cannot occur simultaneously. However, when events in a causal system are exclusive, the normative properties of the whole system change substantially. Are adults sensitive to the consequences of exclusivity for causal reasoning? Here, we systematically manipulated common-effect causal systems to have either exclusive or non-exclusive causes while holding all other factors constant. Adults showed a rich understanding of exclusive systems in making both predictive (Experiment 1) and diagnostic (Experiments 2 and 3) causal inferences. Adults' success in these tasks suggests that exclusivity is an important dimension in human causal reasoning.

Keywords: exclusivity; independence; causal reasoning; predictive; diagnostic

Introduction

Many real-world events exclude the occurrence of others. If it is raining outside, it cannot also be clear, or snowing. If an ion is negatively charged, it cannot also be neutral, or positively charged. Indeed, the world is replete with such mutually exclusive events, and they appear in most complex causal systems, impacting everything from our everyday planning to our scientific theories.

The exclusivity of events also has consequences for the causal properties of the system to which they belong. Exclusivity, by definition, licenses the inference from the presence of one event to the absence of the other (e.g., if it is raining, then it is not sunny). However, exclusivity also informs inferences about causally related events. For instance, both rain and sun cause pedestrians to carry umbrellas (to keep dry or shaded, respectively). In this causal system, two exclusive causes bring about a common effect.

In such a system, exclusive causes are subject to different normative rules than non-exclusive, independent causes[1]. If the causal parameters of a system (i.e., causal powers and base-rates) are held constant, exclusive generative causes will yield greater effect likelihoods than non-exclusive generative causes (see Equations 1 vs. 2). Different calculations are also required when reasoning from effects to causes. For non-exclusive causal systems, the presence of an effect will always enhance the causes' likelihood. The same is not true for exclusive causes (see Equations 3 vs 4). Instead, the presence of an effect will *reduce* the likelihood of any weaker cause's occurrence. In principle, then, exclusivity plays a critical role in causal reasoning.

Despite this clear theoretical distinction between exclusive and non-exclusive causal systems, the latter has received substantially less attention. Much of the research on multi-cause causal systems has focused on independent causes: that is, causes that are appropriately modeled as i.i.d. random variables (for a review, see Rottman & Hastie, 2014). Indeed, causal independence has become a common assumption in models of causal reasoning (e.g., Cheng, 1997; Glymour, 2001; Griffiths & Tenenbaum, 2005).

However, recent work has challenged this assumption in certain circumstances, specifically for modeling mutually exclusive causal events. Fenton et al. (2016) reviewed a range of different approaches to modeling exclusivity in causal Bayes nets, finding that each has substantial weaknesses. Most approaches struggle either to guarantee that the events cannot co-occur or to maintain each event's prior probability. To address these issues, Fenton et al. instead proposed a novel solution, featuring a constraint and auxiliary node. Their solution has been successfully applied to several legal scenarios (Fenton, Neal, & Lagnado, 2013; Vlek et al., 2016).

It is an open question, however, to what extent mutual exclusivity plays a role in human causal reasoning. Recently, Meder and Mayrhofer (2017) demonstrated that people reason accurately about exclusivity in the context of disease diagnosis. However, most studies relating to exclusivity have focused on the use of categories in causal reasoning. For instance, Waldmann & Hagmayer (2006) demonstrated that adults use existing exclusive categories to structure learning

[1] Throughout the paper, we contrast exclusivity with non-exclusivity. However, our non-exclusive systems always feature *independent* causes. Independent causes are, by definition, non-exclusive, but it is worth noting that non-exclusive causes need not be independent. Some non-exclusive causes still share some degree of dependence (e.g., age and income). Thus, while we adopt the terms "exclusive" and "non-exclusive" for clarity, our work truly contrasts exclusive causal systems with independent causal systems.

of a new causal system—even when doing so results in sub-optimal causal predictions. However, while research into the intersection of categories and causality often employs exclusive categories, few studies have directly assessed the specific contribution of exclusivity to people's causal reasoning.

The most promising research on exclusivity comes from studies of category-based induction. Inductive inferences are often interpreted to be a special case of causal reasoning (e.g., Rehder & Hastie, 2001). Early work conducted by Murphy and Ross (1994, 1996) suggested that adults struggle to make inductive inferences about individuals who might belong to any one of several, mutually exclusive categories (e.g., a painting that could be either a Monet or a Renoir). Adults' inferences tended to reflect only the most likely category, rather than reflecting the consideration of information from all possible categories. This finding has been extended to include both artificial and natural stimuli across a range of paradigms (e.g., Malt, Ross, & Murphy, 1995). However, recent work reveals several situations in which adults do consider multiple, exclusive categories. Hayes and Newell (2009) found that adults could successfully integrate information from exclusive categories to make predictive inferences when the cost of neglecting the less likely alternative was made clear, e.g., by associating the less likely category with greater rewards or more serious negative outcomes. In addition, Chen, Ross, and Murphy (2014) extended this research into the realm of decision-making, asking how uncertain categorization affects binary decisions, rather than probabilistic inductive inferences. Here, adults proved more willing to consider multiple, exclusive categories in their reasoning. Interestingly, however, they did so even under conditions of high certainty—when consideration of multiple categories was non-normative. This suggests their consideration of multiple exclusive causes was not grounded in a deep understanding of the causal system but rather in a tendency to focus on given alternatives. In short, while recent evidence has struck a more optimistic tone about adults' capacity for exclusive category-based reasoning, the evidence remains decidedly mixed.

In the present study, we directly test whether adults reason accurately about exclusive causal systems and whether they can appropriately distinguish them from non-exclusive causal systems. Specifically, we present participants with either exclusive or non-exclusive systems in a classic, probabilistic causal inference task for two-cause, common-effect systems. By holding constant all aspects of the system except for the exclusivity of the two causes, we isolate the contribution of exclusivity to causal reasoning. In Experiment 1, we ask whether adults consider exclusivity in their predictive reasoning. In Experiments 2 and 3, we extend this result by examining the impact of exclusivity on adults' diagnostic reasoning.

Experiment 1

In this study, participants were introduced to either exclusive or non-exclusive common-effect causal systems and asked to make predictive inferences about them, reasoning from cause to effect.

Method

Participants Forty-five participants from Northwestern University participated in this study in exchange for course credit. Four participants failed the check question (see below) and so were excluded from analysis.

Procedure Each participant was assigned to either the Exclusive (n = 21) or Non-exclusive condition (n = 20). In each condition, participants read a description of a novel contraption built by Prof. McNutt (Edwards & Rips, 2013). Both contraptions represented common-effect systems. For example, in the Non-exclusive condition, participants were told the device had three components, A, B, and E, each of which could be either on or off. Critically, participants were also informed that components A and B (the causes) were each independently capable of turning on component E (the effect). However, at a given time, A and B could both be on, both be off, or only one of them could be on (i.e., they were independent). In contrast, participants in the Exclusive condition were informed that the device had two components, A and E, but that A always had either positive or negative polarity—but never both.[2] Importantly, component A's ability to turn on E varied as a function of its polarity (see Figure 1).

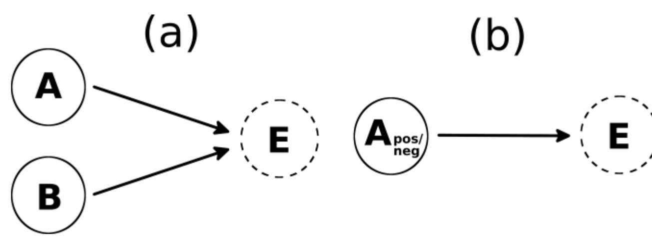

Figure 1: Diagram of the non-exclusive (a) and exclusive (b) causal systems.

Next, we presented all participants with information about the frequency with which each cause turned on (base-rate) and the likelihood that they would turn on component E when on (causal power). This information was held constant across both conditions. The parameters that participants received are given Table 1. In the Exclusive condition, for instance, participants were told the following:

"Here is how the contraption behaves:

[2] Pilot testing indicated that participants had difficulty understanding exclusivity when the exclusive causes were represented as separate components. Thus, to ensure participants understood that the causes were exclusive (could not co-occur), we adopted the already widely known exclusive distinction between positive and negative polarity.

- Part A has positive polarity in 65% of the trials.
- On the trials Part A has positive polarity, it turns on Part E in 80% of those trials.
- Part A has negative polarity in 35% of the trials.
- On the trials Part A has negative polarity, it turns on Part E in 90% of those trials."

Importantly, this describes an *exhaustive* exclusive system. That is, Part A always has either negative or positive polarity; there are no possible alternatives. While it is also possible to reason about non-exhaustive, exclusive causes, we focus on exhaustive causes here for the sake of simplicity.

Table 1: Experiment 1 Exclusive and Non-exclusive System Parameters and Normative Calculations

Cause	Base-Rate	Causal Power	Exclusive Norm	Non-exclusive Norm
A/A_{pos}	.65	.80	.835	.671
B/A_{neg}	.35	.90		

Participants in the Non-exclusive condition were given the same information, except it was attributed to the non-exclusive components (e.g., "Part A turns on for 65% of the trials", "Part B turns on for 35% of the trials.", etc.).

Finally, participants made a predictive inference about the system based on the information they had just read. Specifically, they were asked, "Imagine McNutt runs 100 trials on the contraption. Please indicate the number of times you believe Part E would turn on." Participants provided their answer using a slider with 0 and 100 as endpoints.

We compared the accuracy of people's causal predictions with normative values calculated from the given parameters. For exclusive systems, the likelihood of E turning on can be thought of as the union of two mutually exclusive events in probability theory:

$$(1)\ P(e) = R_{A\_pos}W_{A\_pos} + R_{A\_neg}W_{A\_neg}$$

where R is the base-rate and W is the causal power of the two exclusive causes. The normative likelihood of E is 83.5% (.65*.8 + .35*.9 = .835). The formula for non-exclusive, independent systems is similar but the intersection of these two events (i.e., instances in which both causes turn on E at the same time) is subtracted, under a noisy-OR assumption.

$$(2)\ P(e) = R_A W_A + R_B W_B - (R_A W_A R_B W_B)$$

Thus, the normative likelihood of E turning on in the non-exclusive system is only 67.1% (.65*.8 + .35*.9 – [.65*.8*.35*.9] = .671). If participants are attending to causal exclusivity, predictive likelihoods for E in exclusive causal systems should be greater than non-exclusive causal systems.

At the end of the experiment, participants answered a check question assessing whether they believed the two causes could co-occur or not. Participants who answered incorrectly were excluded from analysis.

Results and Discussion

Participants' effect likelihood ratings suggested they did consider the causes' exclusive or non-exclusive status in making predictive inferences (see Figure 2). Participants in the Exclusive condition (M = 79.5, SD = 5.1) rated the effect as more likely to occur than participants in the Non-exclusive condition (M = 60.4, SD = 22.1), $t(39) = 3.82$, $p < .001$, d = 1.19.

Notably, both groups slightly underestimated the normative likelihood. While the Non-exclusive condition's mean rating did not differ from the normative estimate of 67.1% likelihood, $t(19) = 1.36$, $p = .19$, $d = .30$, participants in the Exclusive condition gave estimates significantly below the normative estimate of 83.5% likelihood, $t(20) = 3.12$, p = .005, $d = .68$. Importantly, participants' predictive inferences were substantially more variable in the Non-exclusive condition, ranging from 16% to 85%, than in the Exclusive condition, which ranged only from 66% to 86%. In addition, participants may simply have been reluctant to give effect likelihoods at near-ceiling levels.

A more fine-grained analysis of participants' responses suggests that far more participants gave a normatively correct response (operationalized as within 2.5 points of the normative answer) in the Exclusive condition (12 participants) than in the Non-exclusive Condition (2 participants), $\chi^2(1) = 10.12$, $p = .001$. That is, over half the participants in the Exclusive condition gave a normative answer, compared to only 10% of those in the Non-exclusive condition. This is likely because calculating the normative answer for the non-exclusive system requires an additional mathematical step (subtracting the causes' intersection). Nevertheless, this disparity emphasizes the facility with which adults reason about exclusive causal systems.

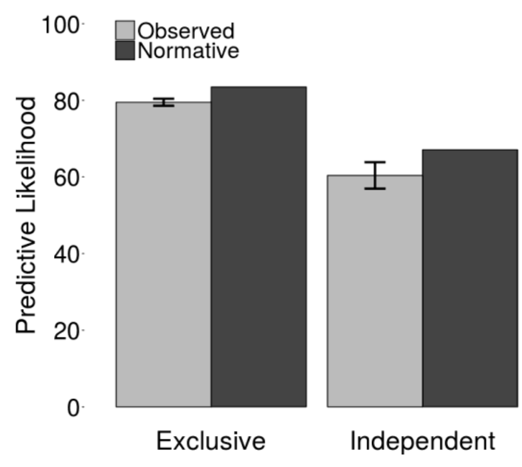

Figure 2: Predictive inference scores from Experiment 1. Error bars show +/- 1 SE.

Experiment 2

While Experiment 1 suggests that adults do consider causal exclusivity when reasoning predictively, it remains an open question whether adults do so when reasoning *diagnostically* as well. Diagnostic reasoning is a common form of causal inference, in which we reason from the presence of an effect to the presence of one of its causes. For instance, we might make inferences about diseases based on symptoms, or the temperature based on people's attire.

The calculations for normative diagnostic inferences, like predictive inferences, differ for exclusive and non-exclusive, independent causal systems. Both formulas use Bayes' rule, but the exclusive system does not consider situations in which both causes occur. This is reflected in the denominator of (3).

$$(3)\ P(A_{pos}|e) = 1 - (1 - R_{A\_pos}) \frac{W_{A\_neg}}{R_{A\_pos}W_{A\_pos} + R_{A\_neg}W_{A\_neg}}$$

For instance, the presence of an abnormal heart rhythm might indicate the presence of either hyperkalemia (excess potassium) or hypokalemia (insufficient potassium)—but certainly not both. On the other hand, for non-exclusive causal systems, the causes might co-occur. For instance, a lawn becoming wet might be due to rain, sprinklers, or both. Such inferences require a different normative rule:

$$(4)\ P(A|e) = 1 - (1 - R_A) \frac{R_B W_B}{R_A W_A + R_B W_B - (R_A W_A R_B W_B)}$$

In Experiment 2, we ask whether adults' diagnostic inferences reflect this normative distinction between exclusive and non-exclusive causal systems.

Method

Participants Forty-seven participants from Northwestern University participated in this study in exchange for course credit. Seven participants failed the check question and so were excluded from analysis.

Procedure As in Experiment 1, participants were randomly assigned to either the Exclusive (n = 20) or the Non-exclusive (n = 20) condition. The descriptions of the contraptions in each condition were identical to Experiment 1. However, the information about each cause's base-rate and causal power was altered slightly to ensure both causes yielded a substantial difference between the normative answers for the exclusive and non-exclusive systems (see Table 2).

Table 2: Experiment 2 Exclusive and Non-exclusive System Parameters and Normative Values

Cause	Base-Rate	Causal Power	Exclusive Norm	Non-exclusive Norm
A/A_{pos}	.40	.80	.372	.529
B/A_{neg}	.60	.90	.628	.814

Next, participants made a diagnostic inference for each cause. For instance, participants in the Non-exclusive condition were asked, "Imagine Prof. McNutt runs the contraption until Part E turns on 100 times. Of these 100 trials where E turned on, how many times do you believe Part [A/B] was on?" Participants again used a slider to record their answer.

For the Non-exclusive condition, the normative answers were 52.9% for Part A and 81.4% for Part B. In contrast, the normative responses in the Exclusive condition were 37.2% (positive polarity) and 62.8% (negative polarity). Thus, the normative difference between exclusive and non-exclusive responses is similar for the weak (Part A/positive polarity) and strong (Part B/negative polarity) causes: 15.7% and 18.6%, respectively.

Finally, participants completed the same check question as Experiment 1, asking if the two causes could co-occur.

Results and Discussion

We submitted participants' diagnostic inferences (see Figure 3) to a mixed 2 (Exclusive vs. Non-exclusive) x 2 (Weak vs. Strong Cause) ANOVA. This analysis revealed a significant effect of condition: participants attributed higher diagnostic likelihoods to causes in the Non-exclusive condition, $F(1,38) = 5.21$, $p = .028$, $\eta^2_G = .11$. As expected, participants also showed a main effect of cause, $F(1,38) = 247.05$, $p < .001$, $\eta^2_G = .45$. However, there was no interaction between cause and condition, $F(1,38) = .32$, $p = .58$, $\eta^2_G = .001$. Indeed, planned comparisons revealed a significant effect of condition for both the weak cause, $t(38) = 2.15$, $p = .04$, $d = .68$, and the strong cause, $t(38) = 2.13$, $p = .04$, $d = .67$.

These ratings show a strong adherence to normative rules (see Table 2). In the Exclusive condition, neither the weak ($M = 38.7$, $SD = 3.6$) nor the strong ($M = 61.7$, $SD = 4.1$) cause's likelihoods differed from the normative response, ps > .05. In the Non-exclusive condition, participants' inferences did not differ from normative levels for the weak cause ($M = 47.9$, $SD = 18.9$), $p > .05$, but were significantly below normative for the strong cause ($M = 69.4$, $SD = 15.5$), $t(19) = 3.47$, $p = .003$, $d = .78$. Thus, participants showed a largely normative pattern of responding, especially in the Exclusive condition.

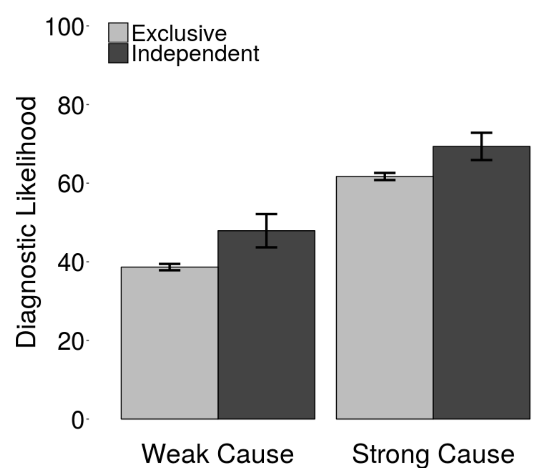

Figure 3: Diagnostic inference scores for both weak (A/A_{pos}) and strong (B/A_{neg}) causes in Experiment 2. Error bars show +/- 1 SE.

However, the normativity of these averaged responses is somewhat deceptive. Only 7 of the 40 participant responses (including both causes) in the Exclusive condition fell into a normative range (+/- 2.5 points) that excluded the cause's base-rates. Moreover, only 1 of the 40 answers in the Non-exclusive condition was in the normative range. This decrease in normative responding relative to Experiment 1, particularly in the Exclusive condition, is likely a result of the heightened mathematical difficulty of the diagnostic calculation. Participants may have become overwhelmed and fallen back on the base-rate as a clear-cut answer.

While participants' individual answers may have been largely non-normative, their diagnostic inferences nevertheless revealed a normative tendency. Participants attributed higher diagnostic likelihoods to non-exclusive causes than to matched exclusive causes, suggesting adults accurately distinguish between exclusive and non-exclusive systems in diagnostic reasoning.

Experiment 3

Normatively, diagnostic reasoning across exclusive and non-exclusive causal systems differs not just in degree but in kind. Experiment 2 suggests that at least on average, adults are sensitive to the difference in degree, giving higher diagnostic inferences for non-exclusive causes. However, when exclusive causal systems are exhaustive, as in Experiments 1 and 2, they always present a difference in kind as well: the diagnostic likelihood of the weakest cause is lower than its base-rate. That is, the occurrence of the effect makes it *less* likely that the weakest cause has occurred whereas in non-exclusive causal systems, the occurrence of the effect will always increase the likelihood of its causes. Importantly, understanding this rule requires little computation, only an understanding of exclusivity and its causal implications.

Therefore, in Experiment 3, we assess whether adults possess an intuitive understanding of the difference between these two systems. We employ the same causal systems as in Experiment 2, but instead of generating a diagnostic inference, participants were asked to compare a cause's diagnostic likelihood with its base-rate.

Method

Participants Forty participants from Northwestern University participated in this study in exchange for course credit. All participants passed the check question.

Procedure The procedure was identical to Experiment 2 with one exception. After learning about the exclusive (n = 20) or the non-exclusive (n = 20) causal system, participants were not asked to provide a diagnostic likelihood. Instead, participants were asked whether the diagnostic likelihood would be higher or lower than the cause's base-rate. For instance, for the weak cause, the base-rate was 40%. As such, participants in the Exclusive condition were asked: "Imagine Prof. McNutt runs the contraption until Part E turns on 100 times. Of these 100 times where Part E turned on, which scenario is more likely:

- Part A turns on *more* than 40 times
- Part A turns on *less* than 40 times"

If adults possess a normative understanding of exclusivity, then they should select the "less" option for the weaker exclusive cause, and the "more" option for the stronger exclusive cause. On the other hand, for the non-exclusive causal system, they should select the "more" option for both causes.

Results and Discussion

Responses were analyzed by condition. In the Exclusive condition, participants selected the "more" option significantly more often for the strong cause than the weak cause, McNemar's $\chi^2(1) = 9.1$, $p = .003$. While 90% of participants believed the strong cause would occur at above-base-rate levels in the presence of the effect, only a minority of participants (35%) believed the weak cause would do so.

In contrast, participants in the Non-exclusive condition did not differ in their beliefs about the weak and strong causes, McNemar's $\chi^2(1) = 1.1$, $p = .29$. In particular, 75% of participants indicated the strong cause would occur at above-base-rate levels, and 50% of participants indicated the weak cause would do so.

Notably, this null result is normative: in the non-exclusive causal system, both causes should increase in likelihood in the presence of an effect. However, it is striking that only half of participants endorsed the weaker cause's enhanced diagnostic likelihood. There are two potential explanations for this finding. First, the rough split in responses may reflect general confusion about the computational steps required to make such inferences. Indeed, the greater complexity for non-exclusive systems is reflected in (4). Alternatively, some people may mistakenly believe that diagnostic inferences for non-exclusive causal systems follow the same principles as exclusive systems. Findings that show preference for singular causal explanations over multivariable explanations (Liljeholm & Cheng, 2007; Lombrozo, 2007) may provide additional evidence for this explanation.

Importantly, adults accurately make the seemingly unintuitive judgment that only a stronger exclusive cause, not a weaker one, is more likely to occur in the presence of its effect. This result suggests that adults possess a rich understanding of exclusivity in causal systems.

General Discussion

The aim of this paper was to assess whether people are capable of distinguishing between exclusive and non-exclusive causal systems. Across three experiments, we examined patterns in people's predictive and diagnostic causal reasoning, looking at both their fine-tuned causal inferences and broad understanding of statistical principle. Experiment 1 demonstrated that participants distinguished between the two systems when reasoning *predictively*, providing greater likelihoods for effects in exclusive causal systems. In addition, Experiments 2 and 3 demonstrated that people's *diagnostic* inferences were also sensitive to the causal implications of exclusive systems. People indicated that causes in exclusive systems were less likely to occur than those in non-exclusive systems, and most recognized that in exclusive systems, weaker causes are actually less likely to occur when the effect is present. In sum, adults accurately distinguish between exclusive and non-exclusive causal systems across a variety of causal reasoning tasks.

Moreover, our findings suggest adults are at least as accurate—if not more accurate—in their reasoning about exclusive systems as non-exclusive systems. Variance in participants' inferences is substantially lower across all experiments for the exclusive systems, and more individual responses fall into the normative range. In Experiment 3, participants' inferences were largely normative for the exclusive system but showed signs of confusion for the non-exclusive system, despite the clear normative answer and the fact that no calculation was required. Thus, adults' causal reasoning may be, if anything, more attuned to exclusive causal systems than non-exclusive causal systems.

Notably, exclusive causal systems can pose a challenge for computational models of causal reasoning, which typically assume causal independence (c.f. Fenton et al., 2016). Our findings emphasize the importance of developing new modeling strategies for incorporating exclusivity, as well as other non-independent causal structures, into causal models.

More broadly, our results illustrate the need for more psychological research on alternative causal structures (c.f., Lucas & Griffiths, 2010; Meder, Mayrhofer, & Waldmann, 2014). Exclusivity is only one of many dimensions which adults may use to reason effectively about causal systems. Theories of human causal reasoning stand to benefit substantially from taking such dimensions into account. We believe the findings presented here are a step in that direction.

Acknowledgments

Thanks to Nathan Couch, Natalie Gallagher, Meghan Salomon-Amend, and Kathy Tian for their comments. This material is based on work supported by a NSF Graduate Research Fellowship under Grant No. DGE-1324585.

References

Chen, S. Y., Ross, B. H., & Murphy, G. L. (2014). Decision making under uncertain categorization. *Frontiers in Psychology*, 5, 1-12.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A general structure for legal arguments about evidence using Bayesian networks. *Cognitive Science*, *37*(1), 61–102.

Fenton, N., Neil, M., Lagnado, D., Marsh, W., Yet, B., & Constantinou, A. (2016). How to model mutually exclusive events based on independent causal pathways in Bayesian network models. *Knowledge-Based Systems*, *113*, 39–50.

Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA, US: The MIT Press.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, 51(4), 334–384.

Hayes, B. K., & Newell, B. R. (2009). Induction with uncertain categories: When do people consider the category alternatives? *Memory & Cognition*, *37*(6), 730–743.

Liljeholm, M., & Cheng, P. W. (2007). When is a cause the "same"? Coherent generalization across contexts. *Psychological Science*, *18*(11), 1014–1021.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*(3), 232–257.

Lucas, C. G., & Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical Bayesian models. *Cognitive Science*, *34*(1), 113–147.

Malt, B. C., Ross, B. H., & Murphy, G. L. (1995). Predicting features for members of natural categories when categorization is uncertain. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 646–661.

Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*(3), 277–301.

Meder, B., & Mayrhofer, R. (2017). Diagnostic causal reasoning with verbal information. *Cognitive Psychology*, *96*, 54–84.

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology*, *27*(2), 148–193.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: the effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology General*, *130*(3), 323–360.

Rips, L. J., & Edwards, B. J. (2013). Inference and explanation in counterfactual reasoning. *Cognitive Science*, *37*(6), 1107–1135.

Ross, B. H., & Murphy, G. L. (1996). Category-based predictions: Influence of uncertainty and feature associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 736–753.

Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*(1), 109–139.

Vlek, C. S., Prakken, H., Renooij, S., & Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law*, *24*(3), 285–324.

Waldmann, M. R., & Hagmayer, Y. (2006). Categories and causality: The neglected direction. *Cognitive Psychology*, *53*(1), 27–58.