Journal of Experimental Psychology:
Learning, Memory, and Cognition
2013, Vol. 39, No. 5, 1327–1343


0278-7393/13/$12.00 DOI: 10.1037/a0031851

# A Quantitative Causal Model Theory of Conditional Reasoning

Philip M. Fernbach
University of Colorado

Christopher D. Erb
Brown University

The authors propose and test a causal model theory of reasoning about conditional arguments with causal content. According to the theory, the acceptability of modus ponens (MP) and affirming the consequent (AC) reflect the conditional likelihood of causes and effects based on a probabilistic causal model of the scenario being judged. Acceptability of MP is a judgment of causal power, the probability that the antecedent cause is efficacious in bringing about the consequent effect. Acceptability of AC is a judgment of diagnostic strength, the probability of the antecedent cause given the consequent effect. The model proposes that acceptability judgments are derived from a causal Bayesian network with a common effect structure in which the probability of the consequent effect is a function of the antecedent cause, alternative causes, and disabling conditions. In 2 experiments, the model was tested by collecting judgments of the causal parameters of conditionals and using them to derive predictions for MP and AC acceptability using 0 free parameters. To assess the validity of the model, its predictions were fit to the acceptability ratings and compared to the fits of 3 versions of Mental Models Theory. The fits of the causal model theory were superior. Experiment 3 provides direct evidence that people engage in a causal analysis and not a direct calculation of conditional probability when assessing causal conditionals. The causal model theory represents a synthesis across the disparate literatures on deductive, probabilistic, and causal reasoning.

*Keywords:* conditional reasoning, causal models, Bayesian networks, disabling conditions, alternative causes



When reasoning about deductive arguments, people are biased to accept conclusions consistent with their beliefs and reject those that are inconsistent, regardless of argument validity (Evans, 2007). In a set of seminal articles, Cummins (1995; Cummins, Lubart, Alksnis, & Rist, 1991) showed that these belief biases follow systematic principles when people reason about conditional arguments with causal content. In the studies, people judged the validity of four argument schemata: Modus Ponens (MP), Modus Tollens (MT), Denying the Antecedent (DA) and Affirming the Consequent (AC), though we focus on just MP and AC in this article. Despite MP being deductively valid and AC being invalid regardless of content, Cummins (1995; Cummins et al., 1991) predicted that for arguments where the antecedent is a cause of the consequent, acceptance rates for MP would be affected by the number of disabling conditions, while AC would be affected by the number of alternative causes for the effect.

In the case of MP, thinking of a disabling condition provides a counterexample to the argument and hence may lead people to reject it. Examples are given below. Cummins (1995; Cummins et al., 1991) predicted that (a) would be judged more acceptable than (b) because the conditional in (a) has fewer disablers; reasons why one could put fertilizer on plants and not have them grow quickly are more available than reasons why one could jump into a pool and not get wet.

(a) If Mary jumped into the swimming pool then she got wet.
Mary jumped into the swimming pool.
Therefore, she got wet.

(b) If fertilizer was put on the plants then they grew quickly.
Fertilizer was put on the plants.
Therefore, they grew quickly.

In the case of AC, alternative causes provide an alternative explanation for the effect and hence make the antecedent seem less necessary. For example, Cummins (1995; Cummins et al., 1991) predicted that (c) would be judged more acceptable than (d). It is hard to think of alternative causes for a gun firing besides the trigger being pulled, but it is relatively easy to think of causes of wetness besides jumping into a swimming pool.

(c) If the trigger was pulled then the gun fired.
The gun fired.
Therefore, the trigger was pulled.

(d) If Mary jumped into the swimming pool then she got wet.
Mary got wet.
Therefore, she had jumped into the swimming pool.

To test these ideas, Cummins (1995; Cummins et al., 1991) asked one group of participants to generate alternative causes and disabling conditions for a set of conditionals and then divided the conditionals into four groups of four conditionals each based on

This article was published Online First April 8, 2013.

Philip M. Fernbach, Leeds School of Business, University of Colorado; Christopher D. Erb, Department of Cognitive, Linguistic, and Psychological Sciences, Brown University.

Experiment 1 was published in the 2010 proceedings of the Cognitive Science Society (Fernbach & Darlow, 2010). This work was supported by a Galner Dissertation Fellowship and an American Psychological Association Dissertation Research Award to Philip M. Fernbach. We are especially grateful to Adam Darlow for many contributions to this work and to Denise Cummins for digging up her data from 1995. Denise Cummins and Mike Oaksford provided excellent comments throughout the review process. Steve Sloman, David Over, Dinos Hadjichristidis, Ben Rottman, and John Lynch also provided helpful feedback on the project.

Correspondence concerning this article should be addressed to Philip M. Fernbach, Leeds School of Business, University of Colorado, 419 UCB, Boulder, Colorado 80309-0419. E-mail: philip.fernbach@gmail.com

the number of alternatives and disablers (many alternatives, many disablers; many alternatives, few disablers; few alternatives, many disablers; few alternatives, few disablers; see Appendix A). A different group of participants was given the arguments based on the 16 conditionals and was asked to judge the extent to which the conclusion could be drawn from the premise. The results supported both predictions; the acceptability of MP was influenced by number of disablers while the number of alternative causes influenced the acceptability of AC. This pattern of results has been replicated many times (e.g., Markovits & Potvin, 2001; Thompson, 1995) and has been found to generalize to more applied settings involving consumers' acceptability ratings of product claims (Chandon & Janiszewski, 2009).

## Prior Theories of Causal Conditional Reasoning

Cummins (1995; Cummins et al., 1991) argued that as the number of alternatives and disablers increases, participants are more likely to retrieve one of them, reducing the perceived necessity (in the case of alternatives) or sufficiency (in the case of disablers) of the causal relations, thereby leading to lower acceptability ratings. The goal of these articles was to show that causal conditional reasoning cannot be reduced to a truth-functional relationship, not to provide a computational model of judgment. Still, Cummins's (1995; Cummins et al., 1991) analysis suggests a potential model, one in which the likelihood of rejecting an argument is related to the number of counterexamples.

In the causal conditional reasoning literature, theories based on the number of counterexamples in memory have been associated with Mental Models Theory (MMT; Johnson-Laird, 1983; Johnson-Laird, Byrne, & Schaeken, 1992). This theory proposes that reasoners represent the content of an argument in a form similar to that of an entry in a truth table, with each entry signifying a possibility (Evans, 1993; Johnson-Laird & Byrne, 2002; Markovits & Barrouillet, 2002). For example, consider the argument from above concerning fertilizer and plant growth. According to mental models theory, an initial model of the conditional is generated in which the antecedent and the consequent are true (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999):

| | |
|---|---|
| Fertilizer on plants | Plants grow quickly |
| . . . | |

In this *initial* model, not all possibilities are represented. The ellipsis above represents an implicit model that could be made fully explicit but does not have any explicit content of its own. If the models were to be made fully explicit through deliberation, the following *full* model would result:

| | |
|---|---|
| Fertilizer on plants | Plants grow quickly |
| ¬Fertilizer on plants | Plants grow quickly |
| ¬Fertilizer on plants | ¬Plants grow quickly |

Above we have the three possibilities that are consistent with the truth of the conditional statement. However, given that the antecedent (Fertilizer on plants) was stated in the argument above, only possibilities in which the antecedent is true will be represented when the MP argument is evaluated, resulting in the following:

| | |
|---|---|
| Fertilizer on plants | Plants grow quickly |

Reasoners may then generate counterexamples to the argument based upon contextual or stored information (Johnson-Laird, 1994). As these counterexamples are generated, further possibilities are represented:

| | | |
|---|---|---|
| Fertilizer on plants | | Plants grow quickly |
| Fertilizer on plants | Too dark | ¬Plants grow quickly |
| Fertilizer on plants | Dry soil | ¬Plants grow quickly |

An analogous process would occur for AC arguments, where alternative causes are generated instead of disabling conditions. The critical question for generating a quantitative prediction based on MMT is how to map this mental representation to a judgment of acceptability. We propose three plausible models based on our reading of the MMT literature and based on how MMT has been interpreted in the causal conditional literature. The most common way to interpret MMT in the causal conditional literature is to claim that acceptability is linearly related to the number of counterexamples (Geiger & Oberauer, 2007; Verschueren, Schaeken, & d'Ydewalle, 2005), and some evidence for this model is provided by De Neys, Schaeken, and d'Ydewalle (2003).

> Model 1: Number of counterexamples; Acceptability of MP decreases linearly with the number of disabling conditions. Acceptability of AC decreases linearly with the number of alternative causes.

A different way to calculate argument strength from mental models is proposed by Johnson-Laird (1994). This model calculates argument strength by evaluating the proportion of possibilities in which the conclusion is true given all of the possibilities that are consistent with the argument. Assuming that each possibility is equally weighted ("equiprobability"), the strength of the above MP argument would be 1/3 given that only one of the three possibilities consistent with the argument represents the conclusion as true.

Notice that this value is different for AC arguments, assuming the fully explicit model. Consider again argument (d):

> If Mary jumped into the swimming pool then she got wet.
> Mary got wet.
> Therefore, she had jumped into the swimming pool.

In this case, two possibilities from the fully explicit model are consistent with the premise of the argument (Got wet) and only one of them represents the antecedent (Jumped in pool) as true:

| | |
|---|---|
| Jumped in pool | Got wet |
| ¬Jumped in pool | Got wet |
| ¬Jumped in pool | ¬Got wet |

Therefore, when no counterexamples are generated, the strength of AC is 1/2, whereas the strength of MP is 1. When, for example, two counterexamples are generated the strength of AC is 1/4, as opposed to 1/3 in MP.

| | | |
|---|---|---|
| Jumped in pool | | Got wet |
| ¬Jumped in pool | | Got wet |
| ¬Jumped in pool | It rains | Got wet |
| ¬Jumped in pool | Water balloon | Got wet |

Therefore, this construal of mental models theory predicts a difference between MP and AC judgments that feature the same number of counterexamples. This model predicts that AC acceptability will decrease nonlinearly with the number of counterexamples.

> Model 2: Equiprobable MMT; Argument acceptability is equal to the number of possibilities in which both the antecedent and the consequent are true divided by the total number of possibilities; MP acceptability = 1/number of disabling conditions + 1; AC acceptability = 1/number of alternative causes + 2.

Proponents of MMT have also proposed theories in which the equiprobability assumption is relaxed. First, Johnson-Laird (1994) argued that reasoners may also take into account the ease with which the possibilities are generated in determining their weight in the argument strength calculations. Consistent with this, in the causal conditional reasoning literature some have argued that certain types of counterexamples are more readily generated than others and that individual differences exist in the propensity to generate certain types of counterexamples (Verschueren, De Neys, Schaeken, & d'Ydewalle, 2002; Verschueren, Schaeken, De Neys, & d'Ydewalle, 2004). More generally, a reasoner may not regard each mental model as equally probable. Thus according to a nonequiprobable MMT, the reasoner can assign probabilities to the possibilities (Johnson-Laird et al., 1999). However, it is difficult to determine how probabilities should be assigned in instances like the above example. As discussed by Geiger and Oberauer (2007), it is unclear how such probabilistic information could be incorporated in a way that allows mental models theory to remain distinct from competing theories. For example, if a reasoner can simply assign a probability to the possibility of interest (e.g., Fertilizer on plants, Plants grow quickly), this would allow a direct route for evaluating the argument's strength that does not require anything particular to the mental models approach.

One way to avoid this concern is by hypothesizing that reasoners do not directly assign a probability to the conditional under evaluation in the argument. Rather, the probability could be inferred after counterexamples have been generated with corresponding probability estimates. For example,

| | | |
|---|---|---|
| Fertilizer on plants | | Plants grow quickly |
| Fertilizer on plants | Too dark (.05) | ¬Plants grow quickly |
| Fertilizer on plants | Dry soil (.05) | ¬Plants grow quickly |

After having generated counterexamples with probability estimates, participants could arrive at an estimate of the conditional under evaluation in the argument. MMT would have little predictive power if there were no constraints on how the probabilities are assigned to counterexamples. Moreover, probabilities could be assigned that yield identical predictions to the other theories discussed below. We therefore assume that probabilities are assigned to counterexamples based on their base rates. For instance, a reasoner might believe that dry soil occurs in 5 out of 100 cases and therefore assign it a probability of .05. Assuming that counterexamples are treated as occurring independently, in the example above this would result in an estimate of .9.[1] To our knowledge, this model has not been previously proposed by proponents of MMT, but we believe it is a plausible instantiation of the theory, and therefore we test it in Experiment 2 by collecting judgments of the base rates of disabling conditions.

> Model 3: Sum of base-rates MMT; Argument acceptability is equal to 1 minus the sum of base rates of counterexamples.

An alternative to mental models theory is the conditional probability theory of conditional reasoning (Liu, Lo, & Wu, 1996; Oaksford, Chater, & Larkin, 2000). On this account, MP acceptability reflects the conditional probability of the consequent given the antecedent, and AC acceptability reflects the conditional probability of the antecedent given the consequent. In support of this idea, Geiger and Oberauer (2007) showed that estimates of the frequency of exceptional cases were a better predictor of acceptability than the number of counterexamples.

Some have argued that generation of counterexamples and judgment of conditional probability or conditional frequency are separate cognitive processes (e.g., Geiger & Oberauer, 2007; Verschueren et al., 2005). For instance, it might be possible in some cases to estimate conditional probability by calculating the conditional frequency of events in memory without thinking of specific disablers (e.g., counting memories of a gun failing to fire and dividing by the total number of attempts to fire a gun). In most of Geiger and Oberauer's (2007) studies, frequency information is given to participants in pretraining. A weakness of the conditional probability theory when construed this way is that it does not explain how people could generate estimates of conditional probability in cases in which frequency information is unavailable. For example, people are capable of judging the conditional likelihood of events for which they have no experience (e.g., Skov & Sherman, 1986). This does not mean the conditional probability theory is incorrect, just that it is silent on the processes and representations that instantiate the conditional probability calculation.

## Causal Model Theory

Our theory combines ideas from MMT and conditional probability theory. We propose that rather than constituting separate processes, the generation of counterexamples and judgment of conditional likelihood are part of a single inferential process based on a mental model that represents the causal structure of the conditional argument being reasoned about. The specifics of our conditional reasoning model come from a theory of cognition called "causal model theory." A causal model is a mental representation of causal structure that includes both qualitative information (the presence and directionality of causal relations between variables) and quantitative information such as the strength of causal relations and the base rates of variables (Pearl, 2000; Sloman, 2005; Spirtes, Glymour, & Scheines, 1993; Waldmann & Holyoak, 1992). Such a model supports inductive inference be-

[1] In this formulation, the possibilities are represented as being mutually exclusive, though in actuality, multiple counterexamples can occur simultaneously. Although this model violates the assumption that individual alternative causes and disablers are independent of one another, it allows for a reasonable estimate of the probability of at least one counterexample occurring with a simple computation that is in the spirit of how reasoning with mental models is traditionally proposed to occur.

cause the probability of unobserved variables can be computed by conditioning on what is known. According to the causal model theory we propose, when confronted with a conditional argument, people embed the antecedent and consequent in a causal model that includes information necessary to calculate conditional likelihood. This includes generating counterexamples, evaluating their role in the causal structure, and estimating their influence. One virtue of causal model theory is that a generic causal structure can be parameterized differently depending on the scenario being reasoned about, yielding probability estimates that are specific to that scenario; the same basic structure can be used to reason about guns firing or apples falling from trees.

Causal model theory has the potential to advance understanding of conditional reasoning beyond MMT and conditional probability theory because it provides a principled way to connect acceptability judgments to the underlying causal beliefs that give rise to those judgments. MMT does not specify how different counterexamples should contribute to argument strength. Conditional probability theory is silent on where judgments of conditional probability come from. In contrast, causal models provide a representational infrastructure and a set of computational rules that transform underlying causal beliefs into quantitative estimates of conditional likelihood. Aside from these theoretical virtues, a great deal of empirical evidence across many domains has been collected over the last 20 years, suggesting that people rely on causal models to learn and make inductive inferences (see e.g., Cheng, 1997; Fernbach, Darlow, & Sloman, 2011a; Griffiths & Tenenbaum, 2009; Holyoak & Cheng, 2011a; Rehder & Burnett, 2005; Sloman & Lagnado, 2005; Waldmann, 2000).

Causal models have also recently been applied to causal conditional reasoning. Ali, Chater, and Oaksford (2011) compared the qualitative predictions of a causal model theory of conditional reasoning with those of MMT by having participants perform diagnostic and predictive causal reasoning tasks. In order to directly compare the theories, Ali et al. identified cases in which the theories make divergent qualitative predictions. For instance, causal models predict a judgment phenomenon called "explaining away" or discounting (Kelley, 1971; Morris & Larrick, 1995; Sloman, Fernbach, & Ewing, 2012). This phenomenon occurs when the likelihood of a cause to bring about an effect is decreased in the presence of another potential cause. However, in the cases presented by Ali et al., MMT does not predict such patterns of discounting. The results were found to be most consistent with the predictions of the causal model framework, suggesting that reasoners' judgments are qualitatively sensitive to the principles of causal model theory not MMT.

Given the theoretical virtues of causal model theory, its ability to fit data across many domains of uncertain judgment, and its success in making qualitative predictions about conditional reasoning, we consider it a promising candidate theory. In this article, we advance this case by proposing and testing a quantitative causal model theory of causal conditional reasoning. In addition to assessing the fit of this model, we compare its performance to the other models introduced above.

## Causal Model Specification

According to Oaksford et al. (2000), if conditional schemata are interpreted in terms of conditional probability, the acceptability of MP maps onto $P$ (*Effect*|*Cause*) and AC to $P$ (*Cause*|*Effect*). Our causal model theory departs from this mapping in one way: Rather than modeling MP as the conditional probability of the effect given the cause, we model it as the probabilistic causal power of the cause or $P$ (*Effect*|*Cause, No Alternative Causes;* Cheng, 1997).

The notion of causal power that our theory relies on draws on work by Cheng (1997) and Pearl (2000), who developed theories of how people learn relations between variables from contingency data. While there is a tradition of research arguing that learners pay attention to contingency (i.e., the change in the probability of the dependent variable when the independent variable is present), proponents of "power" theories argue that learners are actually trying to infer the causal power of the cause or the likelihood that it is efficacious in bringing about the effect when it is present; causal power can be thought of as a measure of causal sufficiency. Whereas Cheng thinks of causal power as a basic primitive of a causal relation between variables, Pearl's theory allows for hidden variables—like disabling conditions—that mediate the relation between the cause and effect. There is disagreement in the literature as to whether these two views are commensurable (see Cheng & Novick, 2005; Luhmann & Ahn, 2005). In this article, we propose two power models: One treats causal power as a primitive, whereas the other explicitly represents disabling conditions as underlying causal power. These two models are broadly consistent with the views espoused by Cheng and Pearl, respectively.

Normatively, predictive conditional probability must be greater than causal power (or equal to it, when the strength of alternatives is zero) because conditional probability increases with the strength of alternative causes while causal power is insensitive to alternative causes. However, previous work has shown that people are not sensitive to the strength of alternative causes when judging predictive conditional probability, even when alternative causes are relevant; when asked for a judgment of predictive conditional probability, people actually provide estimates of causal power (Fernbach, Darlow, & Sloman, 2010; Fernbach et al., 2011a). Thus, we would expect judgments of conditional probability and the output of our causal model theory to be similar in all cases because they are both mediated by the same causal analysis.

This correspondence between judgments of conditional probability and the output of our model raises the question of how causal model theory can be differentiated from conditional probability theory. That is, how is it possible to show that people rely on causal models and do not just calculate conditional probability without performing a causal analysis? In Experiments 1 and 2, we collect judgments of the underlying causal parameters of the conditionals and combine them according to our model to make predictions about the acceptability of arguments based on these conditionals. Conditional probability theory makes no prediction about the relation between these parameters and acceptability, so if the causal model theory predicts acceptability, conditional probability theory cannot explain why. Experiment 3 provides more direct evidence for the use of causal models by showing that when conditional probability is held constant, the causal status of the antecedent makes a difference to the acceptability of MP.

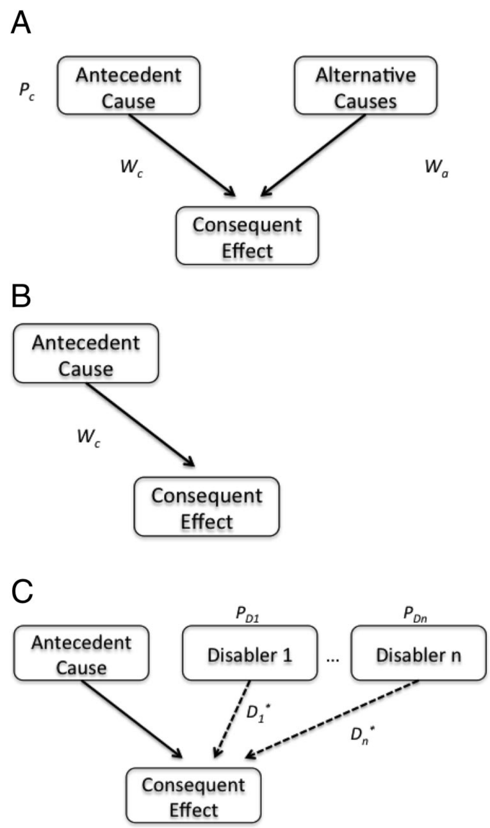


*Figure 1.* Causal models for computing MP and AC acceptability. A: The model for AC. B: The simple model for MP used in Experiment 1. C: The extended model for MP used in Experiment 2. MP = modus ponens; AC = affirming the consequent; $P_c$ = base rate of the cause; $W_a$ = the strength of alternative causes (i.e., the probability of the effect in the absence of the cause); $W_c$ = causal power; $P_D$ = base rate of disabling condition.

## Modeling AC

To derive the model for AC, we assume the "common effect" causal structure in Figure 1A where the antecedent cause and alternative causes can both bring about the effect. The strength of alternative causes is represented by a single parameter, $W_a$ denoting the probability of the effect in the absence of the cause. The antecedent cause is associated with two parameters, its base rate, $P_c$, and its causal power, $W_c$. By assuming that the antecedent cause and the alternatives contribute independently to the effect (i.e., a "noisy-or" model[2]; Cheng, 1997) the expressions in (1) can be derived for AC (for more details see Fernbach et al., 2011a; Waldmann, Cheng, Hagmayer, & Blaisdell, 2008).

$$AC = P(Cause|Effect) = 1 - (1 - P_c)\frac{W_a}{P_cW_c + W_a - P_cW_cW_a}. \tag{1}$$

$W_c$ is the causal power of the cause, the probability that the cause brings about the effect when it is present (e.g., the probability that pulling the trigger causes the gun to fire), $W_a$ is the combined strength of all alternative causes, equivalent to the probability of the effect in the absence of the cause (e.g., the probability of the gun firing given the trigger wasn't pulled), and $P_c$ is the prior probability of the cause (e.g., the probability of the trigger being pulled). AC is a function of all three parameters. It increases with $P_c$ and $W_c$ and decreases with $W_a$.

Diagnostic strength can be thought of as, in part, a comparative measure of the extent to which the antecedent cause is a good explanation for the effect relative to alternative causes. When the strength of alternatives is high, the presence of the effect provides relatively little evidence for the presence of the antecedent cause. Likewise, when the base rate of the antecedent cause is low, more probable alternatives provide a relatively better explanation for the effect.

## Modeling MP

As described above, we model MP in two ways: In the first model (following Fernbach et al., 2011a) individual disablers are not explicitly represented. The structure of the model is shown in Figure 1B. This model assumes that the causal power of the antecedent cause is a primitive. Therefore, MP is modeled simply by this single parameter:

$$MP = P(Effect|Cause, \sim Alternatives) = W_c. \tag{2}$$

In Experiment 1, we assume this model. Participants are simply asked to estimate the causal power of the antecedent to produce the consequent. The model assumes that people have direct access to this parameter. One weakness of this approach in the context of causal conditional reasoning is that this model does not make manifest the connection between disabling conditions and MP acceptability, which is one of our key objectives.

To address this, in Experiment 2, we test an extended model that proposes that beliefs about disabling conditions are used to estimate power. That is, people are not assumed to have direct access to the causal power of a cause for an effect but instead estimate it based on their beliefs about disabling conditions. In this model, the disabling strengths and base rate of individual disablers are the primitives. Thus the model assumes that causal power is fully determined by the disabling conditions, which are combined to yield an over disabling strength, that is, the likelihood that the antecedent cause (when it is present) will fail to bring about the consequent effect.

The structure for this model is shown in Figure 1C. The dashed lines denote the fact that the causal relations between the disablers and the effect are preventative; they serve to reduce the probability of the effect. Each disabler has some likelihood of being present ($P_d$) and when present has some disabling strength $d^*$, a likelihood of successfully stopping the antecedent cause from bringing about the effect. Thus the disabling probability of the $i^{th}$ disabler ($A_i$) is

$$A_i = P_{di}d_i^*. \tag{3}$$

Individual disablers, each with their own disabling strength and prior probability, are combined using the inclusion/exclusion principle to yield an overall probability that the antecedent cause will fail to bring about the consequent, the aggregate disabling proba-

[2] A noisy-or model is a probabilistic generalization of a Boolean exclusive-or. In this model, each cause provides an independent contribution to the likelihood of the effect.

bility ($A'$). This amounts to adding the disabling probability of each disabler (i.e., its base-rate times its disabling strength) and then subtracting out all of the intersections as in Equation 4:

$$A' = \sum_{i=1}^{n} A_i - \sum_{i,j:i<j} A_i A_j + \sum_{i,j,k:i<j<k} A_i A_j A_k - \ldots + (-1)^{n-1} \prod_{i=1}^{n} Ai. \quad (4)$$

The causal power is the complement of this aggregate disabling probability, as shown in Equation 5. This amounts to assuming that the antecedent cause will necessarily bring about the effect if there are no disablers.

$$MP = P(Effect|Cause, \sim Alternatives) = 1 - A'. \quad (5)$$

## Relation Between the Causal Model Theory and the Number of Counterexamples

Our claim is that when judging MP and AC, people take more into account than a the number of alternatives or disablers, and they do so in a way that conforms to the causal model theory. According to the theory, the determinants of AC acceptability are strength of alternatives, the base rate of the cause, and causal power. The number of alternatives is related to the strength of alternatives in that all else being equal, as the number of alternatives increases so does the probability that they will bring about the effect. Therefore, the model predicts that AC will tend to decrease with number of alternatives (as predicted and found by Cummins, 1995; Cummins et al., 1991). However, number of alternatives is not always a good predictor of strength of alternatives. For instance, a large number of highly improbable or weak alternatives should have less effect on judgment then a single probable, strong alternative. Another difference is that according to the model, the base rate of the cause plays an important role in diagnostic strength; a cause that is very improbable is unlikely to have occurred relative to other more likely causes and is therefore not as good an explanation for the effect.

For MP, causal power will tend to be inversely related to the number of disablers because, all else being equal, as the number of disablers increases, the probability that the cause fails to bring about the effect increases. Thus, the model is consistent with the decrease in MP as number of disablers increases. However, as with alternatives, not all disablers are equally likely or equally effective in preventing the effect. A single strong disabler could lead to a lower causal power than several weaker disablers, for instance, making number of disablers an imperfect predictor of causal power. This is reflected in the extended model (Equation 4) where aggregate disabling probability depends on the disabling strengths and base rates of the individual disablers.

## Qualitative Support for the Causal Model Theory

Some trends appear in Cummins's (1995) data that are not predicted by the number of disablers and alternatives. One is that acceptability ratings of AC for conditionals with many alternatives and few disablers were lower than those with many alternatives and many disablers. Both groups had many alternatives and thus should have yielded similar AC judgments according to a model based solely on number of counterexamples. This difference was replicated by De Neys, Schaeken, and d'Ydewalle (2002) who found lower AC ratings for all few disabler items compared to many disabler items, using the same conditionals. To explain this effect De Neys et al. (2002) proposed that when there are many disablers, they interfere with memory search for alternatives, leading to the observed difference. A perusal of the individual conditionals suggests an alternative explanation based on the causal model theory. The two groups appear to vary not just in number of disablers but also in some of the factors that the causal model theory says should affect diagnostic judgments. Specifically, the items that obtain low acceptability scores share the property that the antecedent cause is weak or improbable relative to the strength of alternatives (see Table 1). For instance, jumping into a swimming pool is improbable relative to other causes of wetness. Conversely, high ratings obtain for arguments in which the cause is strong and probable relative to alternatives. There may be many alternatives for a car slowing, but braking is likely the dominant cause. Likewise, studying hard is probably the strongest cause of doing well on a test. Thus, number of alternatives may be equated across groups, but diagnostic strength appears not to be.

Another trend unexplained by the number of alternatives and disablers is that few alternative conditionals obtained slightly higher MP judgments than many alternative conditionals despite being equated across number of disablers. Again, the causal model theory suggests why this may be so. Several of the many alterna-

Table 1
*Mean Acceptability of AC Arguments for Two Groups of Conditionals From Cummins's (1995) Experiment 1*

| Conditional | Acceptability of AC (−3 to 3) |
|---|---|
| Many alternatives, many disablers | |
| If fertilizer was put on the plants, then they grew quickly. | 1.00 |
| If the brake was depressed, then the car slowed down. | 1.00 |
| If John studied hard, then he did well on the test. | 1.50 |
| If Jenny turned on the air conditioner, then she felt cool. | 1.08 |
| Many alternatives, few disablers | |
| If Alvin read without his glasses, then he got a headache. | 0.75 |
| If Mary jumped into the swimming pool, then she got wet. | 0.25 |
| If the apples were ripe, then they fell from the tree. | 1.00 |
| If water was poured on the campfire, then the fire went out. | −0.08 |

*Note*. AC = affirming the consequent.

tive items seem to have somewhat low causal powers (e.g., 'if the apples were ripe then they fell from the tree') while virtually all of the few alternative items have very high causal powers (e.g., 'if the gong was struck then it sounded'). Thus, while number of disablers was equated across groups, causal power may have varied leading to differing MP judgments.

## Experiment 1

To test whether the causal model theory accounts for causal conditional acceptability ratings, we collected judgments of the relevant parameters: the prior probability of the cause ($P_c$), the causal power of the cause ($W_c$) and the strength of alternatives ($W_a$) for Cummins's (1995) conditionals. Using these judgments, we derived predictions with zero free parameters to which we compared Cummins's acceptability ratings.[3] Another implication of our argument is that judgments of the conditional probability of effects and causes should be similar to acceptability ratings and should also be accounted for by the causal model. Thus, we collected predictive and diagnostic conditional probability judgments from a second group of participants.

### Method

**Participants.** 133 Brown University students were approached on campus and participated voluntarily or participated through the psychology research pool in return for class credit.

**Design, materials, and procedure.** All experimental conditions used questions based on the 16 conditionals from Cummins's (1995) Experiment 1 (see Appendix A). We therefore adopted Cummins's (1995) 2 (number of alternatives; few/many) × 2 (number of disablers; few/many) design with four conditionals in each condition. Judgments were on a 0 (*impossible*) to 100 (*definite*) scale.

Seventeen participants provided judgments of the prior probabilities ($P_c$) and strength of alternatives ($W_a$) for the 16 conditionals. The questions were split onto two pages with all of the $P_c$ questions on the first page and all of the $W_a$ questions on the second page. The order of questions was randomized on each page. For each question, we first stated the conditional and then asked the relevant likelihood question. Examples of $P_c$ and $W_a$ questions are given in (e) and (f), respectively.

(e) If John studied hard then he did well on the test.
How likely is it that John studied hard?

(f) If John studied hard then he did well on the test.
John did not study hard. How likely is it he did well on the test?

Two participants interpreted the conditional statement in the $P_c$ questions as indicating that the cause was present and therefore gave ratings of 100 for all of the $P_c$ questions. We removed these responses from the dataset for all subsequent analyses. An additional 21 participants judged causal power ($W_c$). Methods were identical except that there was just one page of questions. An example of a $W_c$ question is given in (g).

(g) How likely is it that John studying hard for the test causes him to do well?

Ninety-five participants provided predictive and diagnostic likelihood judgments, fully within-participants. Each of these participants therefore answered 32 questions, one predictive and one diagnostic for each conditional. Examples of predictive and diagnostic questions are given in (h) and (i):

(h) John studied hard. How likely is it that he did well on the test?

(i) John did well on the test. How likely is it that he studied hard?

This part of the experiment was administered on a computer in the lab. For each question, participants input their answer using the number keys and hit "return" to move to the next question. The order of questions was randomly determined for each participant.

### Results

As expected, $W_a$ was judged higher for many alternative items compared to few alternative items, $t(16) = 13.4$, $p < .001$, and didn't vary across few and many disablers, $t(16) = 1.4$, *ns*. $W_c$ also varied across the number of alternatives manipulation; $W_c$ was judged higher for few alternative items ($M = 83.4$) compared to many alternative items ($M = 73.9$), $t(20) = 4.8$, $p < .001$. This was not intended by Cummins (1995) but confirmed our intuitions about the unexplained trend in MP; weak alternative items seemed to have lower causal powers despite being equated across number of disablers. Surprisingly, $W_c$ did not vary across the many/few disablers manipulation $t(20) = 1.2$, *ns*, suggesting that number of disablers and causal power were not as closely linked as we expected. The low correlation between number of disablers and $W_c$ ($r = -0.11$, *ns*) also supported this conclusion. $P_c$ did not vary across either manipulation.

**Applying the causal model.** Simply computing Equations 1 and 2 using item means would have been inappropriate because the parameter judgments were collected between participants. We therefore used a sampling procedure to generate model predictions. For each conditional we took 10,000 samples each of $W_a$, $P_c$ and $W_c$ uniformly and randomly from participant responses, and calculated Equations 1 and 2 for each set of samples. We therefore generated 10,000 samples of each probability for each conditional and then took the mean over samples for each conditional as the output of the model. Reruns of the model yielded only negligible differences.

**Fits to AC and diagnostic judgments.** Figure 2A depicts Cummins's (1995) acceptability ratings for AC on the X-axis plotted against causal model fits (Equation 1) on the Y-axis for each of the 16 conditionals, along with the least squares regression line. Figure 2B shows diagnostic judgments plotted against model fits. The model predictions were highly correlated with both Cummins's acceptability ratings, $r = .87$, $p < .001$ and the diagnostic judgments, $r = .93$, $p < .001$. To test whether the causal model is a better predictor of AC and diagnostic judgments than the competitor models we first performed separate linear regressions at-

[3] Mean acceptability ratings, number of disablers, and number of alternatives for each conditional from Cummins (1995) Experiment 1 were obtained from Denise Cummins via personal communication in March 2009.

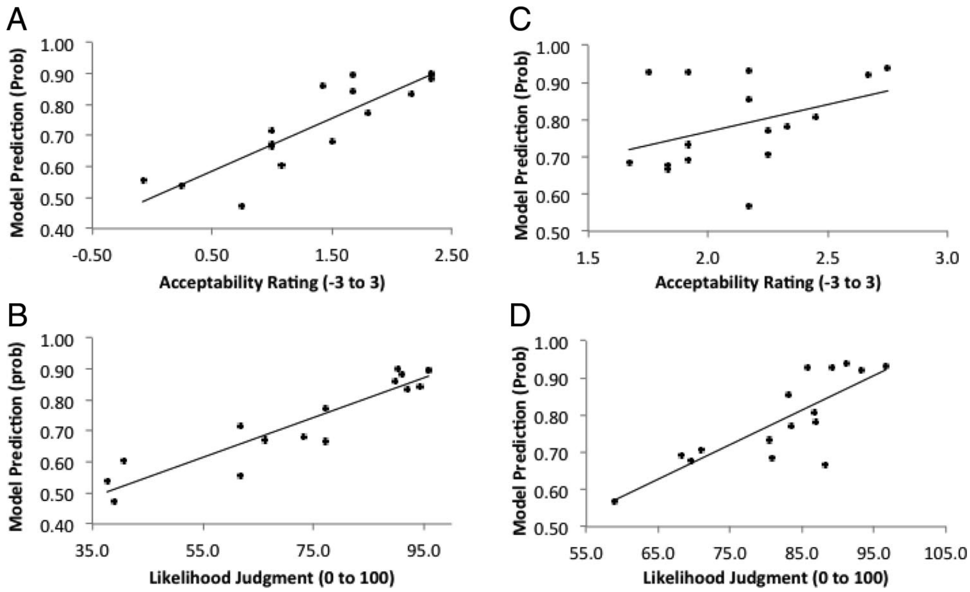


*Figure 2.* Results of model fitting for Experiment 1. A: Causal model fits against Cummins's (1995) AC acceptability ratings. B: Causal model fits against diagnostic likelihood judgments. C: Causal model fits against Cummins's MP acceptability ratings. D: Causal model fits against predictive likelihood judgments. MP = modus ponens; AC = affirming the consequent; Prob = probability.

tempting to predict AC acceptability and diagnostic judgments using the causal model and the first two competitor models (number of alternatives and the equiprobable MMT). The regression results are shown in Table 2. The causal model theory yielded superior fits than the two competitor models.

Next, we performed hierarchical multiple regression analyses using the number of alternatives (the best fitting competitor model) and the causal model predictions as predictors of AC and diagnostic judgments. The causal model accounted for a significant amount of unique variance in AC beyond what number of alternatives accounted for. The model with just number of alternatives explained 55% of the variance in AC. With the addition of the causal model prediction the regression explained 75% of the variance in AC and this change was significantly different from zero, $F(1, 13) = 10.88$, $p < .01$. When the predictors were entered in the opposite order, there was no significant change in variance explained, $F(1, 13) < 1$, *ns*. The pattern was similar for the regression on judgments of diagnostic probability (D). The regression with just number of alternatives explained 52% of the variance in D. Adding the causal model prediction improved the fit to 88% of the variance explained, $F(1, 13) = 37.1$, $p < .001$.

**Fits to MP and predictive judgments.** Figure 2C depicts Cummins's (1995) acceptability ratings for MP plotted against model fits (equal to $W_c$ according to Equation 2). Figure 2D shows predictive judgments (P) plotted against model fits. The pattern of

Table 2
*Regression Results for the Causal Model Theory and Two Competitor Models From Experiment 1*

| Model | Dependent variable | $r$ | $r^2$ | $F$ | Sig. |
|---|---|---|---|---|---|
| Causal model | AC | .87 | .75 | 42.68 | <.001 |
| Number of alternatives | AC | .74 | .55 | 17.39 | .001 |
| Equiprobable MMT | AC | .68 | .46 | 11.84 | .004 |
| Causal model | D | .93 | .87 | 94.75 | <.001 |
| Number of alternatives | D | .72 | .52 | 15.32 | .002 |
| Equiprobable MMT | D | .62 | .38 | 8.60 | .01 |
| Causal model | MP | .39 | .16 | 2.57 | .13 |
| Number of alternatives | MP | .53 | .28 | 5.39 | .036 |
| Equiprobable MMT | MP | .50 | .25 | 4.62 | .050 |
| Causal model | P | .81 | .65 | 25.96 | <.001 |
| Number of alternatives | P | .053 | .003 | 0.039 | .85 |
| Equiprobable MMT | P | .11 | .012 | 0.17 | .69 |

*Note.* AC = affirming the consequent; MP = modus ponens; MMT = mental models theory; D = diagnostic probability judgments, that is, P(cause|Effect); P = predictive judgments; Sig. = significance.

results was less conclusive than for AC. The correlation of number of disablers to MP was just significant, $r = -0.53$, $p = .035$. The correlation of the model fit to MP was just short of significance, $r = .39$, $p = .13$. Predictive judgments were highly correlated with the model, $r = .81$, $p < .001$, but not with number of disablers, $r = -0.04$, *ns* or the equiprobable model, $r = .11$, *ns.* Regression results for the three models are shown in Table 2.

The number of disablers explained a small amount of additional variance in MP when added to a hierarchical regression with just the causal model fit, $F(1, 13) = 4.98$, $p = .041$. When the predictors were entered in the opposite order, change in variance explained was not significant, $F(1, 13) = 2.41$, $p = .15$. The predictive judgments showed a very different pattern. The causal model explained a large amount of additional variance in P when added to a hierarchical regression with just the equiprobability MMT (the best competitor model), $F(1, 13) = 24.61$, $p < .001$, and the equiprobability model did not change the variance explained when the predictors were added in the opposite order.

## Discussion

The results of Experiment 1 were clear for AC judgments. The causal model theory predicted Cummins's (1995) AC acceptability ratings and the diagnostic probability judgments. Number of alternative causes also correlated with both types of judgments but the model captured a large amount of additional variance. The model fitting used zero free parameters making it clear that the better fit was due to participants' sensitivity to the model parameters. These results support our hypothesis that people are sensitive not just to the number of alternative causes, but to diagnostic strength as prescribed by the causal model theory.

The results for MP were less conclusive. The causal power parameter did not fit the MP judgments that closely and was slightly inferior to the number of disablers, though both measures were relatively poor predictors of MP, compared to the fits to AC. Causal power was highly correlated with judgments of predictive probability however. Experiment 2 was intended to strengthen the method in order to provide more conclusive evidence. One weakness of the method in Experiment 1 is that participants may not have had a clear or unitary interpretation of the causal power question (see Fernbach et al., 2010). Thus in Experiment 2 we adopted the extended model for causal power that represents individual disablers in terms of their disabling strengths and base rates. To collect these primitives we asked people to generate disablers for each conditional and then estimate their base rates and disabling strengths. Secondly, we wanted to collect judgments of acceptability and the model parameters within the same experiment as opposed to using model parameters from our participant population to fit acceptability ratings from Cummins's (1995) participants. Finally, by collecting judgments of the base rates of disabling conditions, we were able to test the third competitor model, the sum of base-rates MMT.

# Experiment 2

One group of participants generated disablers for each conditional and estimated the disabling strength and base rate of each disabler. Using these parameters, we computed Equations 3–5 to predict MP acceptability ratings, again with no free parameters. Another group of participants judged the acceptability of MP for the 16 conditionals. We then compared the fit of the model predictions of MMT. The causal model theory predicts that the model should account for more variance in MP judgments than the competitor models.

## Method

**Participants.** Twenty-eight residents of the United States were recruited using Amazon Mechanical Turk (MTurk), and participants completed the disabler generation and rating portion of the experiment. They participated online and received a small payment for completing the study. An additional 14 members of the Brown University community were recruited to judge the acceptability of MP for the 16 conditionals.[4] They received a small payment or course credit and participated in the lab.

**Design, materials, and procedure.** As in Experiment 1, the materials used were based on the 16 conditionals from Cummins's (1995), and the methods were adapted in order to approximate the methods used in that experiment as closely as possible. The disabler generation and rating portion of the experiment was completed online. For each of the 16 conditionals, participants were presented with a situation that was described in two sentences. The first sentence presented the conditional (e.g., If the ignition key is turned, then the car starts). The second stated that the antecedent occurred but the consequent did not (e.g., The ignition key was turned, but the car did not start). Participants were instructed to write down circumstances (disablers) that could make the given situation possible (e.g., the participant could state that the battery was dead). Next, they were asked to imagine 1,000 different instances of the antecedent occurring (e.g., 1,000 instances of ignition keys being turned) and to provide estimates of the base rates and disabling strengths of each disabler: (a) How many times out of these 1,000 instances would the circumstance occur? (b) Of the instances in which the circumstance occurs, how many times would the circumstance cause the consequent not to occur? Item (a) represents the disabler's base rate, and (b), its disabling strength. In case of order effects, half of participants received the conditionals in reverse order.

To reduce confusion, participants were presented with an example problem before beginning the experiment. The instructions and this example problem are shown in Appendix B. Following Cummins (1995), participants were instructed to not dwell on generating disablers and were asked to spend no more than 1 min on that portion of each question. Additionally, participants were instructed to not list slight variations of the same response multiple times. Example responses were provided to illustrate what types of responses were and were not appropriate. Finally, participants were asked whether they felt confident enough in their understanding of the task instructions to continue with the survey and were then given the choice to exit or continue the survey.

[4] An additional 16 participants provided judgments of predictive conditional probability for the 16 conditionals. We also fit those responses with the model and with the number of disablers. The pattern was similar to the fits of the conditional judgments, with the causal model providing a significantly better fit. For the sake of brevity, we do not report those results in the text.

Participants who were assigned to judge the acceptability of MP were seated at a computer in the lab, and participants read instructions adapted from Cummins (1995). They were told that they would see a series of statements and that they would decide whether or not to accept the conclusion that follows the statements. After reading the instructions, they were presented with a modus ponens argument from one of the conditionals. Beneath the conditional argument was a sentence that read, "Given this statement and this fact, please choose the response that best reflects your decision about the conclusion." There was a 7-point response scale with the endpoints labeled "*very sure that I CANNOT draw this conclusion*" and "*very sure that I CAN draw this conclusion*." After answering, they moved on to the next argument until they had judged MP for all 16 conditionals. Presentation order of the 16 arguments was randomized for each participant.

## Results and Model Fits

**Parameter and MP results.** As expected, for the conditionals that Cummins (1995) categorized as having many disablers, participants generated more disablers than for the few disabler conditionals, $Ms = 2.97$ vs. 2.26, $t(27) = 7.5$, $p < .001$. Also corroborating Cummins's results, the many disabler conditionals yielded lower MP acceptability judgments (1–7 scale), $Ms = 6.04$ vs. 6.38, $t(13) = -2.8$, $p < .05$. Two participants were apparently using the logical rule, as they gave the maximal judgment for all of the arguments.

**Model fits.** For each participant, we calculated both the number of disablers for each conditional and the model prediction for causal power for each conditional. To calculate the model prediction for a given conditional, we combined the disabling strengths and prior probabilities for each generated disabler using the function given in Equation 4 and calculated causal power with Equation 5. We then took the average causal power over participants for each conditional as the causal model prediction for that item.

We also generated predictions for the competitor models. The first model predicts that acceptability is linearly related to number of disablers. The second model, the equiprobable MMT, predicts that acceptability is related to one divided by the number of disablers plus one. The final competitor model is the Sum of Base-Rates MMT, where the probability of each disabler is given by its base rate, and acceptability is modeled as one minus the sum of these base rates. To generate this prediction, we used the average sum of base rates over participants for each conditional.

Of the 28 participants, three participants' average implied causal power judgments (based on the causal model theory) were more than three standard deviations outside the mean. For these three participants, parameter estimates were such that predicted MP values were close to zero. In other words, they gave estimates for the base rates and powers of disabling conditions that were much higher than the other participants. We excluded these participants from the main analysis, though we also describe the results with these participants included.

We conducted linear regressions for each model, attempting to predict acceptability ratings. The results are shown in Table 3. As predicted, the causal model theory was superior to the three mental models competitors, $rs = .66$ vs. .41 for the best competitor, the equiprobable MMT. The causal model theory is the only model that explained significant variance, though the competitor models were just short of significance.

As in Experiment 1, we also conducted hierarchical regression to test whether the causal model prediction captured additional variance over the best competitor model, in this case the equiprobable MMT. Adding the causal model prediction to the hierarchical regression including the competitor model as a predictor of MP did indeed explain significant additional variance. Adding the causal model prediction increased the explained variance from 17% to 45%, $F(1, 13) = 6.72$, $p = .02$. When the predictors were added in the opposite order, the equiprobable MMT did not change the explained variance, $F(1, 13) < 1$, *ns.*

Including the outliers in the analysis decreased the variance explained by the causal model theory, though the fit remained significant, $p = .042$ (see Table 3). The fits of the number of disablers and equiprobable MMT models also decreased, but very slightly. The fit of the sum of base-rates model decreased more substantially. The advantage of the causal model theory over the best competitor model, assessed with hierarchical regression, became marginally significant, $p = .08$. To summarize, when the outliers were included, the evidence in favor of the causal model theory was weaker, but still present. The causal model theory was still the only model that explained significant variance and marginally increased the variance explained when added to the best competitor model.

## Discussion

The causal model theory accounted for significant additional variance in judgments of MP compared to the competitor models. This supports our contention that MP judgments reflect causal power, and in turn, the disabling strengths and base rates of disablers. Once again, the model relied on zero free parameters, making the interpretation of the model fit transparent. The method

Table 3
*Regression Results for the Causal Model Theory and Three Competitor Models From Experiment 2*

| Model | Dependent variable | $r$ | $r^2$ | $F$ | Sig. |
|---|---|---|---|---|---|
| Causal model | MP | .66 | .44 | 10.98 | .005 |
| Causal model (including outliers) | MP | .52 | .27 | 5.04 | .041 |
| Number of disablers | MP | .40 | .16 | 2.71 | .12 |
| Equiprobable MMT | MP | .41 | .17 | 2.89 | .11 |
| Sum of base-rates MMT | MP | .37 | .14 | 2.22 | .14 |

*Note*. MP = modus ponens; MMT = mental models theory; Sig. = significance.

was superior to Experiment 1 in that we used the extended model of causal power and obtained judgments of the base rates and disabling strengths of disablers, thus avoiding the issue that participants may not have had a clear or unitary interpretation of the causal power question from Experiment 1.

The variance captured by the causal model theory was still less than for the fits to AC in Experiment 1. We suspect this is because participants may sometimes reflexively apply the logical rule and deem the argument acceptable regardless of content. Evidence for this hypothesis comes from the fact that two participants gave the maximal judgment for each argument, and overall, judgments tended to be quite high and not vary as much as AC judgments. Previous research also supports this interpretation (Reverberi, Pischedda, Burigo, & Cherubini, 2012). Still, the point remains that the causal model theory captured a large amount of variance, suggesting that most MP acceptability judgments are sensitive to the causal model parameters.

One potential criticism of Experiments 1 and 2 is that neither experiment clearly distinguishes causal model theory from the conditional probability theory. The conditional probability theory makes no predictions about the quantitative link between causal parameters and judgments of acceptability, so causal model theory does provide additional predictive and explanatory power for the results of these experiments beyond the conditional probability theory. Still, conditional probability theory has the virtues of being simpler and more general. The objective of Experiment 3 was to provide more direct evidence that people engage in a causal analysis and not merely a conditional probability calculation when evaluating causal conditionals.

## Experiment 3

As suggested by Weidenfeld, Oberauer, and Hornig (2005), one way to provide evidence for a causal analysis process is by comparing responses to causal and noncausal conditionals. Noncausal conditionals are those in which the antecedent is causally unrelated to the consequent. According to our model, since MP is modeled as a judgment of causal power, causal and noncausal conditionals are always predicted to yield different responses. However, this is not always true for the conditional probability model. It is possible to come up with cases where a causal and noncausal conditional have the same conditional probability. Take for example the conditionals in (a) and (b):

(a) If it is over 80 degrees then a marathon runner will sweat during the marathon.

(b) If the marathon is on a Tuesday then a marathon runner will sweat during the marathon.

Conditional (a) is a causal conditional and conditional (b) is a noncausal conditional (since the fact that the marathon is on a Tuesday has no bearing on whether a runner will sweat). In both cases however, it is virtually guaranteed that a marathon runner will sweat. Thus according to the conditional probability model, MP arguments based on both conditionals should be highly acceptable. For example, if the conditional probability model for MP were instantiated as a calculation of conditional frequency in memory (e.g., Geiger & Oberauer, 2007), a participant would think about instances of marathons when it is over 80 degrees and that are run on Tuesday, respectively, and compute the proportion in which a runner sweats. Since marathon runners always sweat, both should lead to similar conditional frequency judgments.

The causal model theory makes a different prediction. For causal conditionals, acceptability should depend on causal strength and thus should not be maximal, unless causal strength is maximal. In contrast, noncausal conditionals should be treated differently because causal power cannot be used as a basis for judgment. Therefore, we predicted that judgments would significantly differ between causal and noncausal conditionals even when conditional probability was equated.

The intuition that causal and noncausal conditionals are cognitively different also inspired the investigation by Weidenfeld et al. (2005). The authors found support for an integrated model of how people interpret and reason about conditionals with causal—but not arbitrary—content. This finding is consistent with our contention that causal and noncausal conditionals are reasoned about in different ways. However, the content of these conditionals and the method of investigation differed in key ways from the present experiment. Weidenfeld et al. presented participants with pseudonaturalistic cover stories that described the relation between two items (e.g., a newly discovered allergic disease in dogs and a substance called Xathylen) as being either causal or noncausal. Additionally, the cover stories that were designed to encourage a noncausal interpretation of the material did so with only limited success. Control questions demonstrated that participants formed noncausal interpretations of these conditionals roughly 53% of the time.

In the present study, we used a methodology that was more similar to those of Experiments 1 and 2. Participants were simply asked to provide estimates or acceptability judgments about familiar events. In addition to maintaining continuity with the previous experiments, we hoped that this method would allow for a more straightforward approach to studying how people reason about conditionals featuring causal versus noncausal content.

Another consideration in designing the study was how to elicit judgments of conditional probability, particularly for noncausal conditionals. Previous work has shown that weak positive evidence can actually drive judgments of predictive conditional probability judgments below the marginal probability of the effect because people focus on the weak evidence at the expense of alternative causes (Fernbach, Darlow, & Sloman, 2011b). Normatively, the conditional probability of an effect given a probability-raising cause must be higher than the marginal probability of the effect. We were concerned therefore that if asked to judge the likelihood of an outcome given an irrelevant antecedent, participants' judgments might not reflect their true beliefs about the conditional probability but would be too low due to the focus on the weakness of the antecedent. To get around this issue, we instead asked people to judge the marginal probability of the consequent without mentioning the antecedent. As long as participants do not see the antecedent as probability lowering, the conditional probability must lie between the marginal probability and one, and since we designed conditionals such that the marginal probability would be equal to or very close to one, the marginal and conditional should be virtually the same.

Table 4
*Conditionals and Results From Experiment 3*

| Consequent | Causal antecedent | Noncausal antecedent | Marginal probability of consequent | Causal power of causal antecedent | Acceptability of MP, causal | Acceptability of MP, noncausal |
|---|---|---|---|---|---|---|
| An American man visits a hospital sometime in the next 10 years. | An American man contracts a flu. | An American man wears black shoes. | 6.40 | 3.93 | 4.68 | 3.51 |
| A person uses the bathroom in the next 48 hr. | A person drinks a can of Coke. | A person plays the drums. | 6.94 | 5.68 | 5.97 | 4.52 |
| An American child speaks fluent English. | An American child's parents read to him. | An American child's parents play tennis with him. | 6.29 | 5.14 | 4.81 | 3.39 |
| A marathon runner sweats during the marathon. | It is over 80 degrees. | The marathon is on a Tuesday. | 6.97 | 5.96 | 6.43 | 5.11 |
| Average | | | 6.54 | 5.18 | 5.47 | 4.16 |

*Note*. All judgments are on a 7-point scale; MP = modus ponens.



## Method

One hundred residents of the United States participated in the main experiment. Sixty-one were recruited from MTurk for a small payment, and 39 were undergraduate business students who participated for course credit. Fourteen additional MTurk participants completed a pretest as described below.

We created causal and noncausal conditionals from four themes for a total of eight conditionals (see Table 4). As in the example above, the conditionals were all designed to have a conditional probability of one regardless of the antecedent. Noncausal antecedents were chosen such that they had no causal relation to the consequent. To verify this, we asked 14 participants to judge whether the antecedent changed the likelihood of the consequent on a 7-point scale with endpoints "*it reduces the likelihood a lot*" and "*it increases the likelihood a lot*" and a midpoint, "*it makes no difference to the likelihood*." As intended, participants judged the noncausal antecedents as making no difference to the consequent (96% of responses; 4% of responses were probability raising) and causal antecedents as being probability raising (80% of responses; 3% of responses were probability lowering, and 17% of responses were "*makes no difference*").

In the main experiment, participants were assigned to one of three conditions. In the "acceptability" condition participants judged the acceptability of Modus Ponens for the eight conditionals using a similar method to Experiment 2. Participants were asked to judge whether to accept the conclusion that the consequent is present following a statement of the conditional and a statement that the antecedent is present. Responses were on a 7-point scale with endpoints, "*very sure I cannot draw this conclusion*" and "*very sure I can draw this conclusion*," and a midpoint, "*cannot tell*." An example is shown below:

> If it is over 80 degrees then a marathon runner will sweat during the marathon.
> It is over 80 degrees.
> Therefore a marathon runner will sweat during the marathon.

In the "marginal probability" condition participants were asked to judge the marginal probability of the four consequents. For instance, one question asked, "How likely is it that a marathon runner will sweat during the marathon?" Responses were on a 7-point scale with endpoints "very sure that the outcome will not occur" and "very sure that the outcome will occur." Finally, in the "causal power" condition, participants judged the causal power of each of the causal antecedents. For instance, they were asked, "It is over 80 degrees. Does it being over 80 degrees cause a marathon runner to sweat during the marathon?" Responses were on a 7-point scale with endpoints "*it will definitely not cause the outcome*" and "*it definitely will cause the outcome*." In all conditions presentation order was randomized for each participant.

## Results

Results for each conditional and overall are shown in Table 4. The results support causal model theory but are inconsistent with conditional probability. As intended, participants saw the marginal probabilities and hence the conditional probabilities for all conditionals as close to maximal ($M = 6.54$). Two of the conditionals yielded marginal probabilities that were slightly below the maximum, so as a robustness check on our results we repeated key tests on each of the conditionals separately.

The first critical prediction that differentiates the theories is that the acceptability of noncausal conditionals should be different from the causal conditionals. This was indeed the case overall, *Mean Difference* = 1.32, $t(36) = 5.19$, $p < .001$, and for each theme separately, all $p$-values $< .01$. In all cases, the noncausal arguments were deemed less acceptable than their causal counterparts.

We also predicted that for causal conditionals, acceptability would be similar to judgments of causal power. This prediction was borne out with all four themes yielding similar judgments between the acceptability of MP for the causal conditionals and their causal power. We submitted these judgments to a repeated-measures analysis of variance (ANOVA) with theme as a within-participants factor and condition (causal power vs. acceptability) as a between-participants factor. There was a significant effect of theme, $F(3, 189) = 17.39$, $p < .001$, suggesting that judgments differed substantially between themes. However there was no effect of condition and no interaction, both $p$-values $> .24$, demonstrating that judgments of causal power and acceptability did not significantly differ. Separate $t$ tests for each theme yielded the same outcome, with no significant differences between causal power judgments and acceptability all $p$-values $> .3$.

For noncausal conditionals, judgments were close to the scale midpoint ("*cannot tell*") indicating uncertainty about the acceptability of the arguments. The average acceptability was not significantly different from the scale midpoint, $M = 4.16$, $t(36) < 1$, *ns*. Analyzed separately, three of the themes yielded nonsignificant differences from the scale midpoint, all $p$ –values $> .14$. One of the themes, "marathon," did exceed the scale midpoint, $M = 5.11$, $t(35) = 3.16$, $p < .01$. We are not sure why this item was judged higher than the others, though (as shown above) it was still lower than the causal conditional from the same theme.



### Discussion

The objective of Experiment 3 was to provide direct evidence that when evaluating a causal conditional, people engage in a causal analysis and not merely a computation of conditional probability. We chose to look at MP because it allowed us to differentiate the theories. We created cases in which conditional probability was maximal but the causal status of antecedents varied. The acceptability of MP for causal conditionals was similar to judgments of the causal power of the antecedents but different from judgments of the marginal probability of the consequents and hence different from conditional probability. Moreover, acceptability of MP for noncausal conditionals differed from the causal conditionals, despite the similarity of their conditional probabilities.[5] Acceptability was near the scale midpoint, indicating that participants were uncertain. Taken together, the results suggest that when assessing a causal conditional, people engage in a causal analysis—in the case of MP they assess the causal power of the antecedent. They do not simply calculate conditional probability. In many cases, the causal analysis will yield responses that are identical to conditional probability, but the current experiment allowed us to dissociate the two theories.

One might argue that a version of the probabilistic theory that relies on subjective rather than objective conditional probabilities could account for these results. We believe that this argument begs the question. When subjective conditional probabilities violate the normative requirements of probability theory, then conditional probability theory provides no guidance about where the subjective probabilities come from. If they come from a causal analysis (as seems to be the case) then causal model theory provides novel explanatory and predictive power that conditional probability theory does not provide. We argue therefore that conditional probability, as distinct from causal model theory, is committed to making its predictions on the basis of objective, not subjective, probabilities, and not with regard to causal beliefs.

## General Discussion

We proposed and tested a causal model theory of causal conditional reasoning. According to the theory, MP acceptability reflects a judgment of causal power, the likelihood that an antecedent cause is effective in bringing about a consequent effect. AC acceptability reflects a judgment of diagnostic strength, the conditional likelihood of the antecedent given the consequent. We tested the model across two experiments by collecting judgments of people's underlying causal beliefs and using them to predict acceptability ratings (from Cummins's, 1995, participants in Experiment 1 and ours in Experiment 2). The results of Experiment 1 were conclusive for AC where the causal model theory captured substantial variance beyond the competitor models. Results for MP were inconclusive, possibly due to a lack of clarity about the meaning of the causal power question. In Experiment 2, we improved the method by testing the extended version of the MP model, which explicitly represents individual disablers in terms of their base rates and disabling strengths. MP acceptability ratings were fit better by the causal model theory than by the competitor models. Taken together, the experiments show that people take into account specific causal beliefs about alternatives and disablers and combine this information in a way that reasonably approximates conditional likelihood based on a causal model. Experiment 3 provided more direct evidence that people assess causal power when judging MP for causal conditionals and do not directly compute conditional probability. For instance, when there is no causal link between consequent and antecedent, people give a judgment reflecting uncertainty, even when conditional probability is high.

We believe that the most important takeaway from these results is that they provide evidence for an intermediate causal modeling process that stands between retrieval of context specific memories and judgment. Since the pioneering work of Tversky and Kahneman (1973) on the availability heuristic, an influential view holds that uncertain judgments reflect a relatively direct link between retrieval and judgment; a hypothesis seems likely when we retrieve many instances of it (or unlikely, in the case of counterexamples). Our results show that MP and AC acceptability judgments cannot reflect a simple counting of the retrieved alternatives and disablers. Instead, judgments depend on information specific to particular alternatives and disablers (base rates, disabling strengths, alternative strengths), and this information is aggregated in a way that approximately respects the dictates of causal model theory. This means that people know (approximately) how the causal role of a variable determines its evidential force. When considering diagnostic strength, for instance, the strength and prevalence of an alternative cause trades off against the base rate of the focal cause. Likewise, when considering causal power, the disabling strengths and base rates of individual disablers determine their contribution to the overall disabling probability. We take our results to suggest that once an alternative or disabler is retrieved from memory it is embedded in a causal structure that determines how information specific to it matters to overall likelihood. This perspective squares with recent proposals suggesting that the plausibility of a causal mechanism is an important ingredient to inference (Cummins, 2010) since people would likely reject an implausible causal link or represent it as having a very low causal power.

We have evaluated three types of models of causal conditional reasoning. All three have some strengths and limitations. Mental Models theory embodies a consideration of alternative

[5] We did not test AC in this experiment, but the causal model theory predicts that AC will increase with the causal power of the cause, which suggests that AC acceptability for causal conditionals will be higher than for noncausal ones. This is the pattern found by Weidenfeld et al. (2005).

causes and disabling conditions, which is indeed central to how people evaluate causal conditionals. MMT also uses simple and psychologically plausible representations and computational rules. However, the theory lacks empirical support. As Ali et al. (2011) have shown, it makes incorrect qualitative predictions. In the current article, none of the instantiations of MMT fit the data as well as the causal model theory we favor. While MMT represents counterexamples, it does not do so in a *structured* way. Evidently, people's representations for judging causal conditionals are more sophisticated than a simple list of possibilities; they also incorporate relations between variables and knowledge about parameters like causal strengths and base rates. In that sense, MMT's computational simplicity also appears to be its biggest limitation.

The probabilistic approach is related to causal model theory in that both theories claim that acceptability maps to conditional likelihood, though they differ in the model they propose for MP. In fact, Ali et al. (2011) saw the two approaches as commensurable. They argued that causal model theory is a specification that explains where the conditional probabilities come from. The current Experiment 3 shows that this is not exactly correct, as MP is better modeled as judgment of causal power than a predictive conditional probability. Nonetheless, we broadly agree with this point.

An important virtue of causal model theory is that it shows how acceptability judgments can be derived from more primitive beliefs in a principled way. For instance, in Experiment 2, we showed that causal power—and hence MP acceptability—can be estimated based on underlying beliefs about disabling conditions. One limitation of this demonstration is that it leaves open the question of where these parameters come from. People may have direct access to them either through learning from direct experience, instruction, analogy, or some other source. In other cases, they might have structured knowledge that goes even deeper, allowing them to estimate the parameters. In that case, the "primitives" in our model would not be psychologically primitive but would reflect the outcome of a causal reasoning process at an even deeper level of knowledge. The depth of knowledge may also differ across domains and across individuals. We see our model as just part of the explanation for how people arrive at judgments of acceptability from more primitive beliefs.

Another ostensible weakness of causal model theory, as compared to the probabilistic approach, is the delimited set of arguments to which it applies. The probabilistic approach is more general in that it applies to all probability judgments, not just causal ones. Some research suggests however that people reason more veridically about probabilities when they are contextualized in terms of a causal model (e.g., Krynski & Tenenbaum, 2007). Thus, causal model theory's specificity may not be a weakness if it can generate predictions about the types of probability judgments that people will struggle with and those they will handle more naturally.

A final limitation is the complexity of the computations required for exact inference with causal models. As the number of variables increases, these computations quickly becoming overwhelming. We doubt that the algorithms people use for inference instantiate the computations exactly. Instead, people probably use simplifying heuristics (Fernbach & Rehder, 2012). These heuristics are evidently quite effective, given the impressive fits to data in Experiments 1 and 2. Future research exploring them would be worthwhile.

A promising development in the cognitive sciences over the last several years has been a synthesis across the historically disparate literatures on deductive and probabilistic inference (Oaksford & Chater, 2001, 2009). Conditional reasoning is a great exemplar of this movement, as the probabilistic approach to conditional reasoning has become a popular and influential view (Over, Hadjichristidis, Evans, Handley, & Sloman, 2007). We see our work (along with Ali et al.'s, 2011) as following in this tradition by synthesizing across another disparate literature, causal reasoning. Our work suggests that we should not think of counterexample generation and conditional likelihood judgment as separate cognitive processes. Instead, counterexamples are an input into a causal reasoning process that has degree of belief as its output. Causal models thus provide a parsimonious account of how people reach uncertain conclusions about causal conditionals.

## References


Ali, N., Chater, N., & Oaksford, M. (2011). The mental representation of causal conditional reasoning: Mental models or causal models. *Cognition, 119,* 403–418. doi:10.1016/j.cognition.2011.02.005

Chandon, E., & Janiszewski, C. (2009). The influence of causal conditional reasoning on the acceptance of product claims. *Journal of Consumer Research, 35,* 1003–1011. doi:10.1086/593292

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104,* 367–405. doi:10.1037/0033-295X.104.2.367

Cheng, P. W., & Novick, L. R. (2005). Constraints and nonconstraints in causal learning: Reply to White (2005) and to Luhmann and Ahn (2005). *Psychological Review, 112,* 694–707. doi:10.1037/0033-295X.112.3.694

Cummins, D. D. (1995). Naïve theories and causal deduction. *Memory & Cognition, 23,* 646–658. doi:10.3758/BF03197265

Cummins, D. D. (2010). How semantic memory processes temper causal inferences. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 207–217). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780199233298.003.0012

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition, 19,* 274–282. doi:10.3758/BF03211151

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Causal conditional reasoning and semantic memory retrieval: A test of the semantic memory framework. *Memory & Cognition, 30,* 908–920. doi:10.3758/BF03195776

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003). Inference suppression and semantic memory retrieval: Every counterexample counts. *Memory & Cognition, 31,* 581–595. doi:10.3758/BF03196099

Evans, J. St. B. T. (1993). The mental model theory of conditional reasoning: Critical appraisal and revision. *Cognition, 48,* 1–20. doi:10.1016/0010-0277(93)90056-2

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment*. London, England: Taylor & Francis.

Fernbach, P. M., & Darlow, A. (2010). Causal conditional reasoning and conditional likelihood. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the Thirty-Second Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning, *Psychological Science, 21,* 329–336. doi:10.1177/0956797610361430

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011a). Asymmetries in predictive and diagnostic reasoning. *Journal of Experimental Psychology: General, 140,* 168–185. doi:10.1037/a0022100

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011b). When good evidence goes bad: The weak evidence effect in judgment and decision-making. *Cognition, 119,* 459–467. doi:10.1016/j.cognition.2011.01.013

Fernbach, P. M., & Rehder, B. (2012). Cognitive shortcuts in causal inference. *Argument & Computation*. Advance online publication. doi:10.1080/19462166.2012.682655

Geiger, S. M., & Oberauer, K. (2007). Reasoning with conditionals: Does every counterexample count? It's frequency that counts. *Memory & Cognition, 35,* 2060–2074. doi:10.3758/BF03192938

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review, 116,* 661–716. doi:10.1037/a0017201

Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process. *Annual Review of Psychology, 62,* 135–163. doi:10.1146/annurev.psych.121208.131634

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1994). Mental models and probabilistic thinking. *Cognition, 50,* 189–209. doi:10.1016/0010-0277(94)90028-0

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109,* 646–678. doi:10.1037/0033-295X.109.4.646

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99,* 418–439. doi:10.1037/0033-295X.99.3.418

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M., & Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62–88. doi:10.1037/0033-295X.106.1.62

Kelley, H. H. (1971). *Attribution in social interaction*. New York, NY: General Learning Press.

Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General, 136,* 430–450. doi:10.1037/0096-3445.136.3.430

Liu, I., Lo, K., & Wu, J. (1996). A probabilistic interpretation of "if–then". *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49,* 828–844.

Luhmann, C. C., & Ahn, W. (2005). The meaning and computation of causal power: A critique of Cheng (1997) and Novick and Cheng (2004). *Psychological Review, 112,* 685–692. doi:10.1037/0033-295X.112.3.685

Markovits, H., & Barouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22,* 5–36. doi:10.1006/drev.2000.0533

Markovits, H., & Potvin, F. (2001). Suppression of valid inferences and knowledge structures: The curious effect of producing alternative antecedents on reasoning with causal conditionals. *Memory & Cognition, 29,* 736–744. doi:10.3758/BF03200476

Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review, 102,* 331–355. doi:10.1037/0033-295X.102.2.331

Oaksford, M., & Chater, N. (2001). The probabilistic approach to human reasoning. *Trends in Cognitive Sciences, 5,* 349–357. doi:10.1016/S1364-6613(00)01699-5

Oaksford, M., & Chater, N. (2009). Précis of "Bayesian rationality: The probabilistic approach to human reasoning." *Behavioral and Brain Sciences, 32,* 69–84. doi:10.1017/S0140525X09000284

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 883–899. doi:10.1037/0278-7393.26.4.883

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology, 54,* 62–97. doi:10.1016/j.cogpsych.2006.05.002

Pearl, J. (2000). *Causality*. New York, NY: Oxford University Press.

Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of object categories. *Cognitive Psychology, 50,* 264–314. doi:10.1016/j.cogpsych.2004.09.002

Reverberi, C., Pischedda, D., Burigo, M., & Cherubini, P. (2012). Deduction without awareness. *Acta Psychologica, 139,* 244–253. doi:10.1016/j.actpsy.2011.09.011

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22,* 93–121. doi:10.1016/0022-1031(86)90031-4

Sloman, S. A. (2005). *Causal models: How people think about the world and its alternatives*. New York, NY: Oxford University Press.

Sloman, S. A., Fernbach, P. M., & Ewing, S. (2012). A causal model of intentionality judgment. *Mind & Language, 27,* 154–180. doi:10.1111/j.1468-0017.2012.01439.x

Sloman, S. A., & Lagnado, D. (2005). Do we "do"? *Cognitive Science, 29,* 5–39. doi:10.1207/s15516709cog2901_2

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. New York, NY: Springer-Verlag. doi:10.1007/978-1-4612-2748-9

Thompson, V. (1995). Conditional reasoning: The necessary and sufficient conditions. *Canadian Journal of Experimental Psychology, 49,* 1–60. doi:10.1037/1196-1961.49.1.1

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232. doi:10.1016/0010-0285(73)90033-9

Verschueren, N., De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Working memory capacity and the nature of generated counterexamples. In W. Gray & C. Schunn (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society. Fairfax,* VA: Cognitive Science Society.

Verschueren, N., Schaeken, W., De Neys, W., & d'Ydewalle, G. (2004). The difference between generating counterexamples and using them during reasoning. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 57,* 1285–1308. doi:10.1080/02724980343000774

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). A dual process specification of causal conditional reasoning. *Thinking & Reasoning, 11,* 239–278. doi:10.1080/13546780442000178

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 53–76. doi:10.1037/0278-7393.26.1.53

Waldmann, M. R., Cheng, P. W., Hagmayer, Y., & Blaisdell, A. P. (2008). Causal learning in rats and humans: A minimal rational model. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind. Prospects for Bayesian cognitive science* (pp. 453–484). Oxford, England: Oxford University Press. doi:10.1093/acprof:oso/9780199216093.003.0020

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. doi:10.1037/0096-3445.121.2.222

Weidenfeld, A., Oberauer, K., & Hornig, R. (2005). Causal and noncausal conditionals: An integrated model of interpretation and reasoning. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 58,* 1479–1513.

(*Appendices follow*)

## Appendix A

The 16 Conditionals Used in Experiments 1 and 2

Many Alternatives, Many Disablers

If fertilizer was put on the plants, then they grew quickly.
If the brake was depressed, then the car slowed down.
If John studied hard, then he did well on the test.
If Jenny turned on the air conditioner, then she felt cool.

Many Alternatives, Few Disablers

If Alvin read without his glasses, then he got a headache.
If Mary jumped into the swimming pool, then she got wet.
If the apples were ripe, then they fell from the tree.
If water was poured on the campfire, then the fire went out.

Few Alternatives, Many Disablers

If the trigger was pulled, then the gun fired.
If the correct switch was flipped, then the porch light went on.
If the ignition key was turned, then the car started.
If the match was struck, then it lit.

Few Alternatives, Few Disablers

If Joe cut his finger, then it bled.
If Larry grasped the glass with his bare hands, then his fingerprints were on it.
If the gong was struck, then it sounded.
If the doorbell was pushed, then it rang.

(*Appendices continue*)

## Appendix B

### Instructions With Sample Problem for the Disabler Generation and Rating Portion of

### Experiment 2:

The problems that you will be answering have two parts. First you will be asked to write down circumstances that could make a situation possible. For example, a question may be something like this:

Please write down as many circumstances as you can that could make the following situation possible.

If you send a package in the mail, then it will reach its destination.
You sent a package in the mail, but it did not reach its destination.

When you are coming up with these circumstances we ask that you keep a few things in mind:
(a) Do not over think these situations. You shouldn't spend more than a minute coming up with answers for this part of the question.
(b) Number your responses. This will make it easier to answer part 2 of the question.
(c) Do not list variations of the same responses multiple times. We are interested in unique circumstances, so if you list multiple responses that are very similar we will consider them a single response.

For instance, consider these example responses to the question stated above:
1. The package was delivered to the wrong address.
2. The package was lost in transit.
3. The package was destroyed in a car accident.
4. The package was destroyed in a plane accident.

Notice that #3 and #4 are very similar, as both refer to the package being destroyed through an accident during shipment. These responses should have been combined into a single response such as "The package was destroyed during shipment."

Next, part 2 asks for an estimate of how many times each of the circumstances listed in part 1 would occur in 1,000 different instances of people sending packages in the mail. For instance, how many times out of 1,000 would a package be delivered to the wrong address? One could estimate that 10 out of 1,000 packages are delivered to the wrong address. Next, the question asks you to estimate how many of these 10 instances of packages being delivered to the wrong address cause the package to not reach its destination. For example, a package that was delivered to the wrong address might get returned to the post office to be delivered to the correct address, so you may estimate that 5 out of these 10 packages actually fail to reach their destination (as is shown in the example below). For other circumstances the package may not be able to reach its destination (for example, when it is destroyed during shipment) and so every time the circumstance occurs the package will not reach its destination (as illustrated in the example responses below).

Here is what part 2 of the example problem would look like:
Imagine 1,000 different instances of people sending packages in the mail.

Provide estimates for the following questions for each of the circumstances stated above:
(a) How many times out of these 1,000 instances would the circumstance occur?
(b) Of the instances in which the circumstance occurs, how many times would the circumstances prevent the package from reaching its destination?

Here are some example responses:
(a)
1. 10 out of 1,000
2. 3 out of 1,000
3. 1 out of 1,000
4. 1 out of 1,000

(b)
1. 5 out of 10
2. 2 out of 3
3. 1 out of 1
4. 1 out of 1

Received April 19, 2012
Revision received January 6, 2013
Accepted January 11, 2013 ■