

## Naive theories and causal deduction

DENISE DELLAROSA CUMMINS  
*University of Arizona, Tucson, Arizona*

Evidence is presented that implicates two factors in deductive reasoning about causality. The factors are alternative causes and disabling conditions (factors that prevent effects from occurring in the presence of viable causes). A causal analysis is presented in which these factors impact on judgments concerning causal necessity and sufficiency, which in turn determine deductive entailment relations. In Experiment 1, these factors were found to impact causal deductive judgments more strongly than did logical form. In Experiment 2, causal deductive judgments were found to vary as a function of familiarity with a particular causal relationship: The more familiar the causal relationship, the less willing reasoners were to accept conclusions based on them.

A growing body of evidence indicates that humans show systematic response preferences when reasoning about classes of events that have high adaptive consequences (Cheng & Holyoak, 1985, 1989; Cosmides, 1989; Cosmides & Tooby, 1989; Cummins, 1995a; Gigerenzer & Hug, 1992; Manktelow & Over, 1991). Recent developmental evidence suggests that these preferences begin as domain-specific, innate constraints and develop into naive theories that are brought to bear during reasoning about domain-relevant events (e.g., Carey, 1986; Leslie, 1987; Spelke, 1991, 1994; Wynn, 1991). The domains that have received the most attention recently are physical and social reasoning. In the domain of physical reasoning, developmental researchers have shown that infants as young as 3 months of age appreciate that objects are solid, rigid, permanent entities whose movements are continuous in space and time (Baillargeon, 1987; Spelke, 1991). In contrast, the appreciation of inertia and gravity emerges much later (around 9 months) and is much less reliably demonstrated (Spelke, 1991). In fact, these characteristic differences continue to appear in naive adult physics reasoning, where errors are rarely observed on trajectory problems unless inertial and gravitational forces must be taken into consideration (McCloskey, 1983; Spelke, 1991, 1994).

Similar emergent patterns are observed in the performance of social reasoning tasks. Within the first few months of life, human infants distinguish between living and nonliving things (Tronick, Als, Adamson, Wise, & Brazelton, 1978) and among a variety of emotional facial expressions (Boccia & Campos, 1989). By the age of

2½, children distinguish between moral and social rules (Smetana, 1981; Turiel, 1983), and between their own and others' internal states (Leslie, 1987). By 3 years of age, they begin to display the social content effect in deductive reasoning; that is, they reason flawlessly about social rules (such as permissions and obligations) but fail miserably on formally identical problems that have nonsocial content (Cummins, 1995b; Girotto, Gilly, Blaye, & Light, 1989; Girotto, Light, & Colbourn, 1988; Light, Blaye, Gilly, & Girotto, 1989; Light, Girotto, & Legrenzi, 1990). The social content effect continues to appear in the naive reasoning of adults. For example, adults rarely make errors when asked to select, for example, which of four bar customers must have their drinks or identification checked in order to ensure that they are not violating the social permission rule "If a customer is drinking beer, then he or she must be at least 21 years of age." Yet they typically fail miserably when asked to solve formally identical problems based on equally familiar but nonsocial content (e.g., "Every time I go to Miami, I travel by car") (Griggs & Cox, 1982). Performance shifts on social as opposed to nonsocial reasoning tasks have also been observed among other species, particularly nonhuman primates (Cheney & Seyfarth, 1985).

Because human reasoning has traditionally been characterized as a content-free, syntactically based process, content effects like these have typically been categorized as reasoning errors, and they have been attributed to the impact of factors outside of the system itself (e.g., "world knowledge") (Braine & O'Brien, 1991; Evans, 1989; Rumin, Connell, & Braine, 1983). Recently, however, several researchers have begun to argue that the human reasoning system is best characterized as a collection of domain-specific modules that are innate (Cosmides, 1989; Cosmides & Tooby, 1989; Cummins, 1995b) or acquired through experience (Cheng & Holyoak, 1985, 1989; Cheng, Holyoak, Nisbett, & Oliver, 1986). These researchers argue that content effects reflect the access and use of these modules whenever a reasoning scenario with a module-relevant content relevant is encountered.

---

This project was supported by BRSO S07RR07002 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health, and by a Social and Behavioral Sciences Small Grant award (University of Arizona). I would like to thank Kelly Johnson for her invaluable assistance in conducting this work, and Alvin Goldman, Karen Wynn, Ken Forster, and Douglas Medin for comments on the paper. Requests for reprints should be sent to D. D. Cummins, Cognitive Science, University of Arizona, Tucson, AZ 85721 (dcummins@ccit.arizona.edu).

Although a good deal of research has been devoted to the investigation of naive physical reasoning and naive deductive social reasoning, surprisingly little has been devoted to an equally fundamental type of reasoning—naive causal deduction. This is particularly surprising given that innate sensitivity to causal events has been observed among human infants (Leslie & Keeble, 1987), young children (Sedlack & Kurtz, 1981; Shultz, 1982), and members of other species (Garcia & Koelling, 1966), and that causal content frequently produces robust content effects on deductive reasoning tasks (see, e.g., Markovits, 1986; Roberge, 1982). Instead, the bulk of research on causal reasoning has been devoted to causal *attribution*, wherein the reasoner must determine which of a number of possible causes in fact produced a particular effect (e.g., Chapman & Chapman, 1969; Cheng & Novick, 1991, 1992; Downing, Sternberg, & Ross, 1985; Kelley, 1973). Precious little research (e.g., Cheng & Nisbett, 1991; Staudenmeyer, 1975) has been devoted to investigating causal *deduction*—that is, *the inferences that reasoners believe are entailed by their causal hypotheses*. Entailment is a deductive relation. Briefly stated, an inference is entailed by a set of beliefs if accepting the beliefs and rejecting the inference would constitute a contradiction. A conclusion is entailed if it *must* be true given that the premises upon which it depends are true.

In a recent paper, my colleagues and I attempted to address this dearth of knowledge concerning naive causal deduction as defined above (Cummins, Lubart, Alksnis, & Rist, 1991). We found that causal deduction was exquisitely sensitive to two factors, which we termed *alternative causes* and *disabling conditions*. An alternative cause is simply a cause (other than the one cited in the

causal rule under consideration) that is capable of evoking the effect cited in the rule. A disabling condition is an event that could prevent an effect from occurring in the presence of a viable cause. For example, consider the causal rule “If the brake was depressed, then the car slowed down.” This rule admits of many alternative causes that could slow the car down (e.g., going uphill, engine trouble, etc.), as well as of many disabling conditions that could prevent the car from slowing down even if the brake were depressed (e.g., brake lines cut, icy road, etc.). In contrast, the causal conditional “If Larry grasped the glass with his bare fingertips, then his fingerprints were on it” admits of few alternative causes for Larry’s fingerprints’ ending up on the glass, and few disabling conditions that would prevent his fingerprints from appearing there.

In this paper, I propose a mechanism to explain how and why these factors influence causal deduction: I argue that they influence the way reasoners represent causal rules by affecting beliefs concerning the causal necessity and sufficiency of a rule’s components. Specifically, alternative causes cast doubt on the *necessity* of the cause in question for bringing about the effect, whereas disabling conditions cast doubt on the *sufficiency* of the cause to bring about the effect. The evaluation of necessity and sufficiency relations produces the entailment relations that characterize naive causal deduction. This is depicted in Figure 1. When a reasoner is told that a particular effect has occurred and is asked to determine whether a particular cause has produced it, the reasoner considers the possibility that other causes might have produced the effect instead. If no other cause is considered possible, then the reasoner deems the cause in question to be necessary for producing the effect and there-

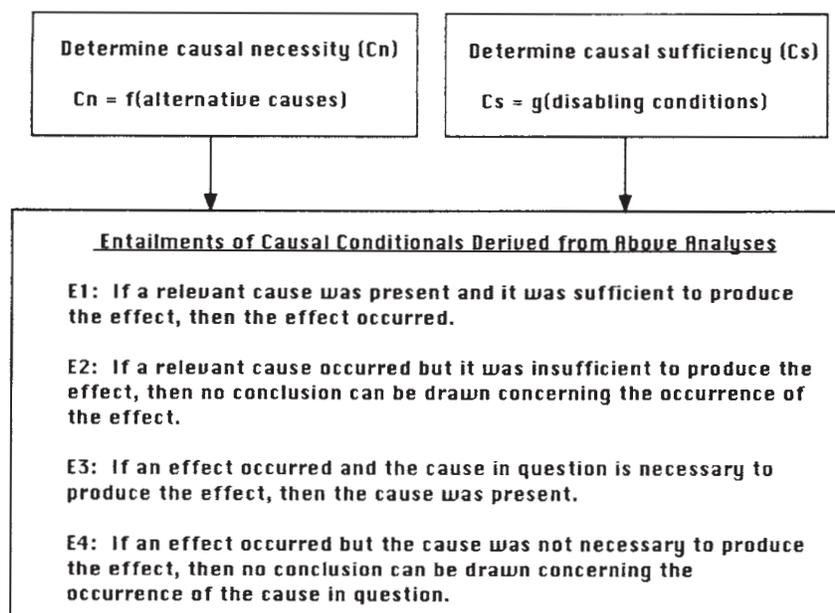


Figure 1. An analysis of naive causal deduction based on evaluations of causal necessity and sufficiency. These evaluations determine entailment relations. See text for details..

fore concludes that the cause *must* have occurred given that the effect is known to have occurred. In contrast, if one or more alternative causes could have produced the effect, then the cause is deemed unnecessary for having produced the effect, and concluding that the cause in question produced the observed effect is not warranted. Similarly, when a reasoner is told that a particular cause occurred, and is asked to determine whether its typical effect occurred as well, the reasoner considers the possibility of disabling conditions—factors that could have prevented the effect from occurring even though a viable cause was present. If none are believed possible, then the occurrence of the cause is deemed sufficient to have produced the effect, and concluding that the effect occurred is entailed. If one or more disabling conditions are believed possible, then the occurrence of the cause is deemed insufficient to have produced the effect, and concluding that the effect will occur is *not* entailed.

Although the results of Cummins et al. (1991) are consistent with this analysis, a stronger test would consist of evaluating the effect of conditional reversal on reasoning. The causal analysis described above predicts that the effects of alternative causes and disabling conditions will remain constant even when the order of the conditional components are reversed—that is, from “If <cause>, then <effect>” to “If <effect>, then <cause>”. In contrast, a formal (truth-functional) analysis predicts no difference due to this reversal. For example, consider the following two arguments:

(a)

If the brake was depressed, then the car slowed down.  
The brake was depressed.  
Therefore, the car slowed down.

(b)

If the car slowed down, then the brake was depressed.  
The brake was depressed. Therefore, the car slowed down.

Formally, (a) is an instance of modus ponens and (b) is an instance of affirming the consequent. If these conditionals are treated as truth-functional conditionals, then truth-functional logic specifies that the antecedent is sufficient (but *not* necessary) for the consequent, and the consequent is necessary (but *not* sufficient) for the antecedent, regardless of the meanings of the antecedent and consequent. Therefore, in reasoning about modus ponens, one is asked to consider whether knowing for certain that the antecedent is true is sufficient to infer that the consequent is true is well. Because the antecedent is sufficient for the consequent of a truth-functional conditional, the answer is yes; modus ponens is a valid argument, and its conclusion can be drawn without fear of producing a contradiction. In contrast, when reasoning about affirming the consequent, one is asked to consider whether knowing for certain that the consequent of the conditional is true is sufficient to

infer that the antecedent is true. Because the consequent is necessary but *not* sufficient for the antecedent, the answer is no; affirming the consequent is *not* a valid argument, and drawing a conclusion is not warranted. If naive human reasoners (those who haven’t had logic training) reason according to the semantics of the truth function for the truth-functional conditional, then they should accept the conclusion of the first argument and reject the conclusion of the second, regardless of familiarity with the causal relationship described. The *syntactic* components of the argument under consideration should account for any variation in performance on the two arguments.

In contrast, a causal analysis based on consideration of alternative causes and disabling conditions yields the opposite prediction. Regardless of the syntactical differences in these arguments, in both cases a conclusion about an *effect* is stated. Therefore, conclusion judgments should be influenced by the consideration of *disabling conditions*—that is, factors that could prevent the car from slowing down even though the brake was depressed. The consideration of disabling conditions casts doubt on the sufficiency of the cause for producing the effect in question; that is, it influences one’s judgment concerning *causal* sufficiency. In this case, numerous factors could disable the cause–effect relationship under consideration (e.g., brake lines cut, or faulty, icy roads, etc.). A causal analysis based on disabling conditions would predict that both conclusions would be rejected, regardless of the syntactic properties of the arguments.

Consider now the following two arguments:

(c)

If the car slowed down, then the brake was depressed.  
The car slowed down. Therefore, the brake was depressed.

(d)

If the brake was depressed, then the car slowed down.  
The car slowed down. Therefore, the brake was depressed.

Again, formally, the first argument (c) is an instance of modus ponens, and the second (d) is an instance of affirming the consequent. So, again, under a formal analysis, the conclusion of (c) should be accepted and that of (d) should be rejected. But, under a causal analysis, regardless of their formal properties, both arguments state a conclusion about a *cause*. Therefore, we would predict that conclusion judgments should be influenced by the consideration of *alternative causes*. Because many other factors could cause a car to slow down other than the brake’s being depressed (e.g., going uphill, engine failure, tire damage, etc.), this casts doubt on the necessity of the cause in question for producing the effect; that is, it casts doubt on causal necessity. Therefore, once again, we would predict that both conclusions would be rejected, regardless of the fact that (c) is a valid argument and (d) is invalid.

Consider now the following arguments:

(e)

If Larry's fingerprints were on the glass, then he grasped it with his bare hands.

Larry's fingerprints were on the glass. Therefore, he grasped the glass with his bare hands.

(f)

If Larry grasped the glass with his bare fingertips, then his fingerprints will be on it.

Larry's fingerprints were on the glass. Therefore, he grasped the glass with his bare fingertips.

Formally, (e) is an instance of *modus ponens* a valid argument, whereas (f) is an instance of affirming the consequent, an invalid argument. By causal analysis, both arguments state a conclusion about a cause, and a causal analysis would evoke consideration of other factors that would cause Larry's fingerprints to be on the glass besides his touching it with his hands. Because few (if any) come to mind, the cause in question should be judged necessary for the effect to occur, and the conclusion should be accepted.

Similar analyses can be offered for *modus tollens* (valid) and denying the antecedent (invalid):

(g) *Modus Tollens*

If the brake was depressed, then the car slowed down. The car did *not* slow down. Therefore, the brake was *not* depressed.

(h) Denying the Antecedent

If the brake was depressed, then the car slowed down. The brake was *not* depressed. Therefore, the car did *not* slow down.

Once again, according to a causal analysis, the important thing is whether one is making a decision concerning a cause or an effect. Accordingly, (g) should be influenced by consideration of alternative causes. Because there are many, reasoners should be highly likely to reject this valid argument. On the other hand, (h) should be influenced by consideration of disabling conditions. Because there are many of these as well, reasoners are likely to reject this invalid argument.

The impact of premise reversal was investigated in Experiment 1. The causal analysis offered here, however, also yields a surprising and counterintuitive prediction: If our deductive causal judgments depend on consideration of alternative causes and disabling conditions, then reasoners should be *more* likely to accept arguments based on unfamiliar causal relationships than arguments based on familiar ones. This prediction was investigated in Experiment 2, in which reasoners were asked to draw conclusions about causal scenarios that differed in familiarity.

## EXPERIMENT 1

The purpose of this experiment was to test the relative contributions of causal and formal properties on causal deduction. It was hypothesized that when an argument based on causal content is detected, our naive theories concerning causal relationships trigger a causal—as opposed to syntactic or formal—analysis of the relationship in question. This means that, regardless of the formal properties of the argument, reasoners become concerned with *causal* necessity and sufficiency—that is, the necessity of a particular cause for bringing about a particular effect, and its sufficiency for bringing it about. Reversing the order of a causal conditional constituent propositions from “if <cause>, then <effect>” to “If <effect>, then <cause>” should produce a reversal in the factors that impact on deductive judgments. A formal analysis predicts no changes in reasoning performance.

To test these predictions, the same manipulation as that used by Cummins et al. (1991) was employed, in which conditionals that described concrete cause–effect relationships were embedded in MP, MT, DA, and AC arguments. The conditionals were chosen on the basis of pilot work in which subjects were required to actually generate alternative causes and disabling conditions for a series of causal conditional statements. Conditionals that varied widely in terms of the average number of items generated were then chosen as materials in the study. Conditionals in the present study that take the form “If <cause>, then <effect>” constitute a replication of Cummins et al. Predictions based on these conditionals are depicted in Figure 2a. When the conditional takes the form “if <cause>, then <effect>,” *modus ponens* and *modus tollens* require the reasoner to make a judgment about whether the effect occurred given that the cause occurred. Therefore, these arguments should be influenced by the number of disabling conditions that characterize the causal relationship in question. Affirming the consequent and denying the antecedent require the reasoner to make a judgment about whether a particular cause in fact produced a known effect. Therefore, these arguments should be influenced by the number of alternative causes that could produce the effect in question.

Figure 2b shows the predictions when the conditional is reversed (i.e., “If <effect>, then <cause>”). A causal analysis predicts a reversal of the factors that influence judgments on these arguments. When the conditional takes this form, *modus ponens* and *modus tollens* require the reasoner to make a judgment about whether a particular cause in fact produced a known effect. Therefore, these arguments should be influenced by the number of alternative causes that could produce the effect in question. Denying the antecedent and affirming the consequent require the reasoner to make a judgment about whether the effect occurred given that the cause occurred. Therefore, these arguments should be influenced by the number of disabling conditions that characterize the causal relationship in question. In other words, a causal analysis predicts that reasoning is influenced by

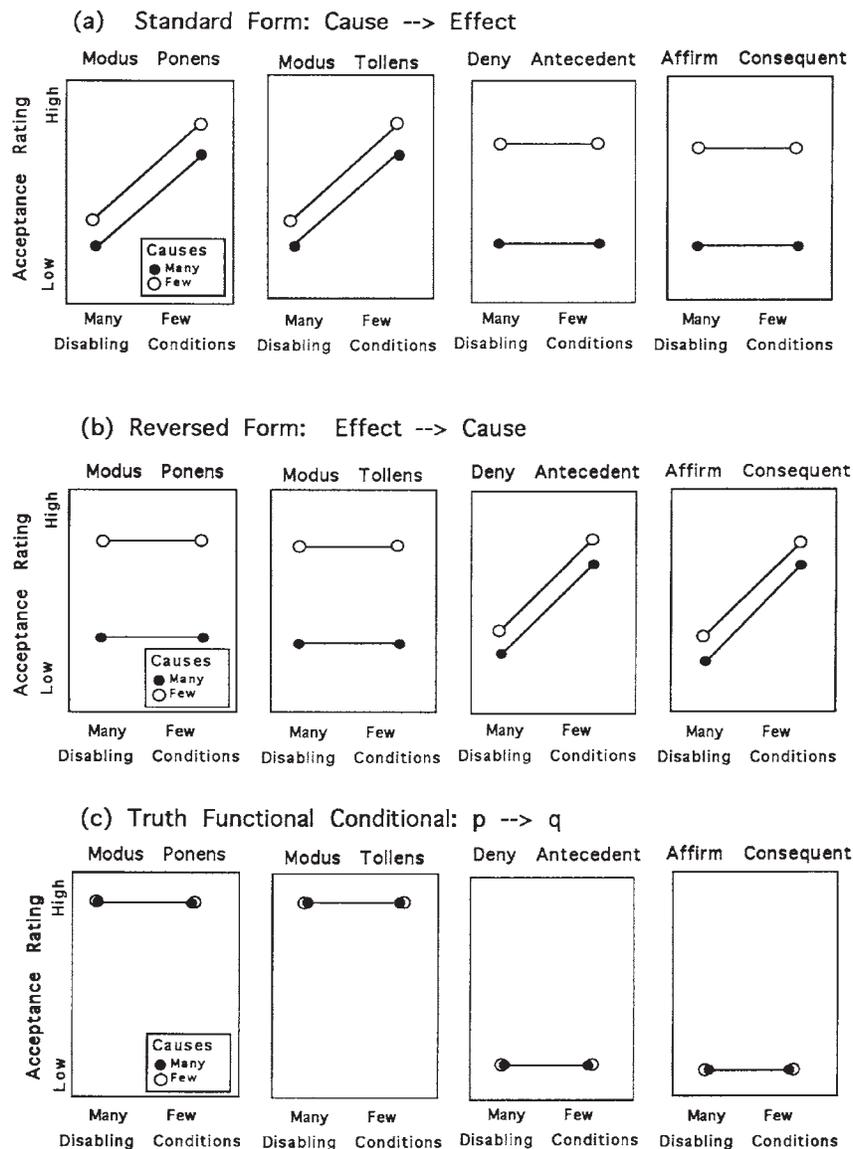


Figure 2. Predictions of acceptability ratings for four types of arguments (modus ponens, modus tollens, denying the antecedent, and affirming the consequent) based on a causal analysis (2a–2b) and a formal analysis (2c). See text for details.

whether a judgment must be made about a cause or an effect, regardless of the form of the argument.

Finally, Figure 2c shows the predictions for a formal analysis. Here, reasoning is influenced solely by argument type. It does not matter whether one must make a judgment about a cause or an effect, nor does it matter whether many or few causes and disabling conditions characterize the causal relationship under consideration. Only argument type matters.

**Method**

**Subjects.** Twenty-four students from introductory psychology courses at the University of Arizona served as subjects in the main experiment. A different set of 80 students enrolled in introductory psychology served as subjects for the generation task that was used

to select materials for the main experiment. All subjects received class credit for participating, and none had had training in logic.

**Generation task materials.** There is much variety among the causal conditionals used in the literature on deduction. Some involve global statements of scientific laws stated in present or future tense (e.g., If water is heated to 100° C at sea level, it will boil), while others describe instances of causation in the past tense (e.g., If the brake was depressed, then the car slowed down). They also vary in terms of whether the constituent propositions refer to classes of objects/people (e.g., trees, Republicans) or to individuals (e.g., Larry, the Rock of Gibraltar). The materials used in Cummins et al. (1991) varied among themselves on these factors as well. The effect of these factors on naive causal deduction is not known, but they affect the normative representation assigned. For example, propositions referring to classes of objects/people are represented as universally quantified statements, while those referring to individuals are treated as simple or existentially quantified statements.

Tense influences normative representations as well. The open question is whether these factors influence consideration of alternative causes and disabling conditions. If they do, a confound would exist between the number of alternatives that appear to characterize a causal conditional and the conditional's syntactic properties, making it impossible to attribute reasoning performance differences unequivocally to the former. For this reason, it was decided to systematically vary these factors on the generation task.

A pool of 52 conditionals was constructed which (1) seemed to describe causal relations and (2) seemed to differ in terms of number of potential alternative causes and disabling conditions on the basis of the intuitions of the experimenter and assistants. One half of these mentioned general classes of objects (e.g., "car," "tree," etc.), and the other half mentioned particular individuals (e.g., "Susan," "Larry"). Past and present tense forms of each were constructed.

For the generation tasks, the conditionals were typed one to a sheet and were bound into booklets in randomized order. Each booklet contained conditionals of the same type (general or individual) and tense (past or present). An example of the *cause generation* task is as follows:

Statement: If Jenny turns on the air conditioner, then she will feel cool.

Fact: Jenny feels cool, but she did *not* turn on the air conditioner.

Please write down as many factors as you can that could make this situation possible.

The instructions for the task stressed the importance of producing items that were reasonably different from each other. An example was given for a situation in which a tree had acorns on the ground beneath it, but the tree was *not* an oak. Sample acceptable items were "someone brought them," or "they were there from a previously growing tree." Sample unacceptable items were "a squirrel brought them," "somebody dropped them," "a young girl dropped them," and "Little Red Riding Hood dropped them for the wolf." It was explained that the unacceptable set would be scored only as one item since it represented simple variations on the theme of the acorns' being dropped or brought by an agent.

An example of the *disabling condition generation* task is the following:

Statement: If Jenny turns on the air conditioner, then she will feel cool.

Fact: Jenny turned on the air conditioner, but she does *not* feel cool.

Please write down as many factors as you can that could make this situation possible.

Once again, the instructions for the task stressed the importance of producing items that were reasonably different from each other. An example was given for a situation in which an oak tree did *not* have acorns on the ground beneath it. Sample acceptable items were "a squirrel took them," "a fire burned them," "the oak is too young," and "the oak is dead." Sample unacceptable items were "a squirrel took them," "somebody gathered them," "a young girl picked them up," and "Little Red Riding Hood gathered them for her grandmother." It was explained that the unacceptable set would be scored only as one item since it represented simple variations on the theme of the acorns' being moved by an agent.

**Generation task procedure.** Subjects were run in groups of 5–10. As mentioned earlier, 26 general class (e.g., "tree," "car") and 26 individual (e.g., "Larry," "Susan") were constructed which seemed (1) to describe causal relations and (2) to differ in terms of number of potential alternative causes and disabling conditions on the basis of the intuitions of the experimenter and assistants. To ensure that our intuitions matched those of our subjects, we required them to first perform a causal rating task on the conditionals and then to actually generate alternatives. For the *causal rating task*, 26 conditionals of the same type (all general or all individual) and

tense (all past or all present) were typed onto a single sheet of paper. The conditionals appeared on the sheet in one of four randomized orders. Above the list of conditionals appeared the following rating scale:

| Strength of Causal Relationship |           |      |          |        |             |         |
|---------------------------------|-----------|------|----------|--------|-------------|---------|
| 0                               | 1         | 2    | 3        | 4      | 5           | 6       |
| None                            | Very Weak | Weak | Moderate | Strong | Very Strong | Perfect |

Subjects were instructed to reserve 0 for conditionals that did not seem to describe causal relationships. As an example, they were shown a permission conditional (i.e., "If you are at least eighteen years old, then you can drink beer"). It was explained that being 18 does not *cause* a person to drink beer. The relationship between age and the permission to drink alcohol is determined by law, not by any causal mechanism.

Instructions for the causal rating task were distributed and were read aloud as the subjects followed along. Time was allowed for questions. They were then given 15 min to complete the task. After that, the sheets were collected and the subjects were instructed to turn over the generation task booklets that were placed face down on their desks.

The generation task design involved a  $2 \times 2 \times 2$  between-groups design, the variables being type of item generated (causes vs. disabling conditions), class of causal event (general vs. individual), and tense of conditional (present vs. past). There were 10 subjects in each cell. The top sheet of the generation task booklet included written instructions for the task, which were read aloud to the subjects. Time was allowed for questions. The subjects were given 1.5 min per conditional for generating alternatives. (In previous pilot work, subjects had generally tended to "run dry" after approximately 1 min.) After that, the subjects were debriefed and dismissed.

**Causal rating and generation task results.** Mean causal strength rating, averaged over 10 subjects for each conditional, was 3.86, with a standard deviation of .66. No conditional was given a causal strength rating of zero. The mean causal rating assigned (3.86) corresponds to a verbal rating close to *strong causal relationship*, corroborating our intuitions that to our reasoners, these causal conditionals described causal relationships.

The generation protocols were scored by two independent raters. Less than 2% of the generated items were disallowed. Interrater reliability for the cause generation task was .93, and for the disabling generation task, .78. Greater disagreement existed in scoring the latter task primarily because of differences in the raters' willingness to collapse items into categories (e.g., "loose connection on distributor cap" and "faulty spark plugs" might be counted as two disabling conditions by one rater and collapsed into the single category of "engine malfunction" by another). Nonetheless, since conditionals were ultimately selected on the basis of *relative* differences in items generated, minor differences like these that produced a moderate interrater reliability score washed out.

The mean number of alternative causes generated and the mean number of disabling conditions generated were computed, averaged over 10 subjects for each conditional. Overall means (and standard deviations) were 3.02 (1.02) for causes generated and 2.75 (.66) for disabling conditions generated.

To investigate the impact of tense and class of causal event on subjects' ability to *generate* items, the mean number of items generated by each subject was computed (by averaging across conditional for each subject). These means were then subjected to an analysis of variance (ANOVA) with three between-subjects variables: task (cause generation vs. disabling condition generation), item type (general class nouns vs. individuals), and tense (past tense vs. present tense). The analysis returned a single significant result, the main effect of task [ $F(1,72) = 7.73$ ,  $MS_e = .31$ ,  $p < .01$ ]. On the average, subjects generated *more alternative causes* (3.02) than disabling conditions (2.75) for the statements. Since the other

two factors washed out, this means that the ability to generate alternative causes or disabling conditions did *not* depend on the tense of the conditional or on whether the conditional referred to a general class of objects or to individuals. The same pattern of results obtained when an item analysis was conducted. This indicates that differences in the number of alternative cause and disabling conditions that characterized these conditionals were due *not* to the conditionals' syntactic/formal properties, but instead to the nature of the causal relationships described by the conditionals. This is an important finding, because the analysis of causal deduction offered here appeals to assessment of causal necessity and sufficiency and not to formal/syntactic properties of conditional premises.

A final analysis was conducted in order to investigate whether the strength of the causal relationship described by a conditional was influenced by syntactic properties. A regression equation was used to predict mean causal strength ratings from the mean number of causes generated, mean number of disabling conditions generated, item type (contrast coding for individual vs. general items) and tense (contrast coding for past vs. present). This equation accounted for less than 1% of the variance ( $R^2 = .009$ ). Therefore, causal strength ratings were *not* affected by any of these factors. This means that regardless of whether the causal relationship referred to individuals or to classes of objects, was expressed in past or present tense, or was characterized by few or many alternative causes or disabling conditions, it was still perceived as a moderate to strong causal relationship. For example, depressing a brake was considered a true cause for slowing down a car even though many things could cause a car to slow down and even though many things could prevent the car from slowing down even when the brake was depressed. The important point of this analysis is that it rules out the alternative explanation that alternative causes or disabling conditions affect causal deductions indirectly by reducing perceived causal strength. Any differences in reasoning performance on the main task must be due to the impact of these factors on the inference process itself.

**Reasoning task materials.** To choose conditionals that varied from each other in terms of mean number of alternative causes and disabling conditions, the items used in the generation tasks were ordered on two lists. On the first, they were ordered from highest to lowest mean alternative causes generated; on the second, they were ordered from highest to lowest mean disabling conditions generated. These orderings ignored whether the item was a "general" or "individual" item, since this factor did not contribute to any variance in the analyses above, thereby providing a large pool from which to choose. Sixteen items were then chosen from these lists, consisting of 4 conditionals that appeared in the top third of both lists (many causes and many disabling conditions), 4 that appeared in the bottom third of both lists (few causes and few disabling conditions), 4 that appeared in the top third of one list and the bottom third of the other (many causes and few disabling conditions), and 4 that appeared in the bottom third of the first list and the top third of the second (few causes and many disabling conditions). These materials therefore constituted a  $2 \times 2$  manipulation of the alternative causes and disabling conditions factors. They are presented in the Appendix along with their respective means.

Materials for the main experiment consisted of arguments based on the causal conditionals presented in the Appendix. These conditionals represented a  $2 \times 2$  manipulation of alternative causes and disabling conditions, as described above. The conditionals were embedded in the four argument types (MP, MT, DA, and AC), producing a total of 64 arguments for each subject to evaluate. In addition, another set of 64 arguments was constructed by using the same conditionals, but with the order of their constituent components reversed (i.e., "If <cause>, then <effect>" was reversed to read "If <effect>, then <cause>").

Each argument was typed on a separate sheet of paper, and below the argument, the following statement and rating scale ap-

peared: "Given this statement and this fact, please circle the response that best reflects your decision about the conclusion."

|  |      |                  |               |                                       |      |              |
|--|------|------------------|---------------|---------------------------------------|------|--------------|
|  |      |                  |               |                                       |      |              |
| Very<br>Sure                           | Sure | Somewhat<br>Sure | Can't<br>Tell | Somewhat<br>Sure                      | Sure | Very<br>Sure |
| That I CANNOT draw this<br>conclusion. |      |                  |               | That I CAN draw this conclu-<br>sion. |      |              |

These pages were bound into booklets in randomized order, with each booklet containing conditionals of only one order (i.e., either standard causal order or reversed order).

**Reasoning task procedure.** Subjects were run in groups of 5–8. They were verbally instructed that they would see a series of statements, and their task was to decide whether or not to accept the conclusions that followed the statements. They were shown a copy of the rating scale and were told to indicate their decisions by placing a mark on the scale that best reflected their decisions. Care was taken to ensure that they understood that making a mark on the left side of the scale indicated that they believed the conclusion could *not* be drawn, making a mark on the right indicated that they believed the conclusion *could* be drawn, and making a mark in the middle indicated that they could not tell one way or the other. The nature of the confidence ratings was also explained. They were told that there were no right or wrong answers. It should be pointed out that the present instructions did not represent an attempt to rapidly school subjects in the meaning of validity as in many studies (and thereby to introduce an instruction confusion element in subject performance); rather, the subjects were encouraged to reason as they would in everyday circumstances. This provides a clearer picture of the inferences that reasoners believe are entailed by causal conditionals.

## Results and Discussion

Rejection probability for all statistical tests reported in this paper was .05, unless otherwise indicated. Significant omnibus  $F$  ratios for interactions were followed by simple effects tests (Keppel, 1973). Tukey's HSD test was used for pairwise comparisons of means, unless specific contrasts are described. Effects involving repeated measures are reported as significant if they pass the Greenhouse and Geisser (1959) correction criterion.

The numbers  $-3$  to  $+3$  were assigned to the rating scale in such a way that  $-3$  represented a judgment of *Very sure cannot draw this conclusion*, and  $+3$  represented the judgment *very sure can draw this conclusion*, with zero representing *can't tell*. Each subject rated arguments based on four different conditionals within each  $2 \times 2 \times 4$  cell, and the mean of these four observations was calculated. These means were then subjected to an ANOVA with conditional order (standard vs. reversed) as a between-groups variable, and argument (MP, MT, DA, AC), number of causes (many vs. few), and number of disabling conditions (many vs. few) as repeated measures.

The main effects of argument, alternative causes, and disabling conditions were significant [ $F(3,66) = 10.82$ ,  $MS_e = .71$ ,  $p < .001$ ;  $F(1,22) = 39.50$ ,  $MS_e = .74$ ,  $p < .001$ ;  $F(1,22) = 9.99$ ,  $MS_e = .41$ ,  $p < .01$ ]. Each of these was modified by higher order interactions with conditional order. Specifically, the three-way interaction of order, argument, and causes was significant [ $F(3,66) = 12.53$ ,  $MS_e = .38$ ,  $p < .001$ ], as was the three-way interaction of

order, argument, and disabling conditions [ $F(3,66) = 8.28, MS_e = .47, p < .001$ ]. These interactions are depicted in Figure 3.

Comparing Figures 2 and 3 makes it apparent that the manipulation of conditional order influenced argument ratings in the way predicted by a causal analysis. When the conditional statement was of the form “If <cause>, then <effect>,” ratings for MP and MT were influenced by the number of disabling conditions that characterized the causal relationship, while ratings for DA and AC were influenced by the number of causes.<sup>1</sup> These results replicate those of Cummins et al. (1991). Also as predicted, reversing the order of the conditional constituents (i.e., “If <effect>, then <cause>”) produced a reversal in the factors that influenced reasoning. With reversal, ratings for MP and MT were influenced by the number of *causes* that characterized the causal relationship, while DA and AC were influenced by the number of *disabling conditions*.<sup>2</sup> Clearly, reversing the order of the conditional statements directly influenced causal deduction. Conclusions about the role a cause played in producing an effect was influenced by the average number of alternative causes that characterized the causal relationship, regardless of the argument’s formal (syntactic) properties. Similarly, conclusions about the occurrence of effects was in-

fluenced by the average number of disabling conditions that characterized the causal relationship, regardless of the argument’s formal properties.

An interesting question is how these conditionals were interpreted (internally represented) by the reasoners. Numerous proposals have been put forth by philosophers of science concerning the proper formal treatment of causal conditionals and causal conditional reasoning (e.g., Goodman, 1955; Lewis, 1973; Mackie, 1974). The upshot of these analyses is that causal conditionals, in many cases, ought not to be represented as simple truth-functional conditionals. As one logic text puts it,

Sentences expressing causal relations represent a connection that is stronger than material implication. An exhaustive analysis of such sentences would take us beyond the bounds of truth-functional logic . . . Therefore, we can treat such conditionals as if they were material implications, even though they are also much more; but we must recognize that, in so treating them, we sacrifice much of their meaning. (Baum, 1989)

The simplest way of investigating subjects’ interpretations with the present data was to collapse the rating scale to two points (*accept* or *don’t accept/can’t tell*) and then analyze the pattern of responses across arguments. A truth-functional biconditional interpretation consists

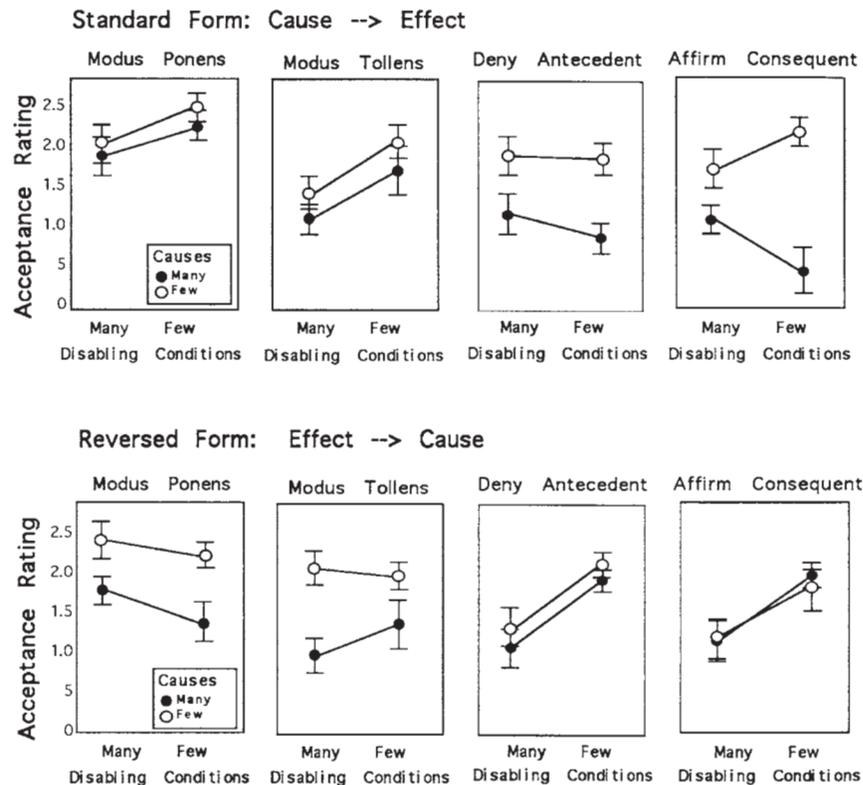


Figure 3. Mean acceptance ratings for arguments based on contextualized, familiar causal conditionals presented in either standard (cause → effect) or reversed (effect → cause) form. The rating scale ranged from -3 (very sure cannot draw this conclusion) to +3 (very sure can draw this conclusion), with 0 representing can’t tell.

of accepting all four arguments; a truth-functional conditional interpretation consists of accepting MP and MT while rejecting DA and AC. Consistent with Staudenmeyer (1975), the results indicated that the overwhelmingly favored interpretation for these causal conditionals was the biconditional. Biconditional interpretations accounted for nearly half (49%) of all response patterns; other patterns—including the truth-functional conditional pattern—each accounted for less than 5% of the overall responses.

More important to the causal analysis offered here, however, is the fact that reasoners' propensity to interpret these causal conditionals as biconditionals varied as a function of alternative causes and disabling conditions, as the following analysis clearly shows: The proportion of biconditional interpretations produced by each subject within each alternative causes  $\times$  disabling condition cell was computed. These proportions were subjected to an ANOVA with conditional form (standard or reversed) as a between-subjects factor and alternative causes (many and few) and disabling conditions (many and few) as repeated measures. The analysis returned two significant effects, the main effect of alternative causes [ $F(1,22) = 16.20$ ,  $MS_e = .05$ ,  $p < .001$ ] and the main effect of disabling conditions [ $F(1,22) = 8.80$ ,  $MS_e = .04$ ,  $p < .01$ ]. The directions of these means indicated that conditionals describing causal relationships for which few alternative causes existed tended to be interpreted as biconditionals more often than conditionals for which many alternative causes existed (59% vs. 40%). Similarly, conditionals with few disabling conditions were interpreted as biconditionals more often than those with many (56% vs. 43%). These factors did not interact with each other, *nor did they interact with the order of the conditional (standard or reversed causal order)*.

To summarize, sensitivity to the possibility of alternative causes and disabling conditions completely dominated these reasoners' judgments, almost to the exclusion of the syntactic properties of the conditionals and the arguments in which they appeared. This result provides strong support for the causal analysis set forth in this paper wherein causal necessity and sufficiency judgments are influenced by the consideration of alternative causes and disabling conditions, and causal inference depends on evaluations of causal necessity and sufficiency.

## EXPERIMENT 2

The results of the previous experiment indicate that part of our naive causal reasoning strategies includes the consideration of alternative causes and disabling conditions. When reasoning about causes, we consider the possibility of alternative causes of an effect; when reasoning about effects, we consider conditions that might prevent the effect from occurring in the presence of a viable cause.

But what happens when we are presented with causal relationships that are unfamiliar to us? A formal analy-

sis would predict that familiarity should have no impact because it is the syntactic properties of the argument that matter. The causal analysis offered here predicts that because no knowledge concerning alternative causes and disabling conditions is available to reasoners in the unfamiliar case, they should be *more* willing to accept arguments based on unfamiliar causal relationships. In other words, because no information exists to cast doubt on the necessity or sufficiency of the cause to bring about the effect, reasoners should be more likely to accept arguments based on unfamiliar causal relationships. Moreover, this willingness to accept should decline to the extent that the unfamiliar relationship bears a similarity to a familiar causal relationship for which many causes and disabling conditions are known. The greater the similarity between the unfamiliar and a known causal relationship, the more likely the reasoner is to import knowledge concerning alternative causes and disabling conditions into the reasoning process.

To test these predictions, the familiarity of the causal content was manipulated in Experiment 2. The four syntactic arguments (MP, MT, DA, AC) were embedded in the following three reasoning scenarios:

*Familiar:* Julie was considering getting up at dawn to go for a long bicycle ride. She hates to ride when the streets are wet, though, and she thought to herself "*If it rains, then the streets will be wet.*"

*Foreign:* Phedra is an alien on a planet that is similar to ours except that the chemical structure of its elements are different. Periodically, a sticky substance called "thardron" forms, and clings to the planet's surface. This is called "thardronning."

Phedra was considering getting up at dawn to go for a long ride on something similar to a bicycle. She hates to ride when the streets are sticky, though, and she thought to herself "*If it thardrons, then the streets will be sticky.*"

*Analogy:* Phedra is an alien on a planet that has a weather system similar to ours except that the chemical structure of its elements are different. Periodically, clouds made up of a substance called "thardron" form, which are similar to clouds of water vapor. Thardron also falls onto the planet in drops similar to our rain, and takes some time to evaporate. The aliens use thardron in many of the same ways we use water.

Phedra was considering getting up at dawn to go for a long ride on something similar to a bicycle. She hates to ride when the streets are wet with thardron, though, and she thought to herself "*If it thardrons, then the streets will be wet.*"

The first causal scenario is one with which most people should be familiar, and it is also one that pilot work indicated is characterized by relatively many alternative causes (for the streets being wet) and disabling conditions (reasons why the streets might be dry even though it rained). Consideration of these factors should reduce the likelihood that reasoners will accept conclusions based on this scenario.

In sharp contrast, the second scenario describes a completely foreign causal relationship. Because of its unfamiliarity, it should be difficult to come up with alternative causes and disabling conditions, making the conclusions based on it appear more plausible.

The third scenario, however, describes a foreign causal relationship that bears a strong similarity to a familiar one (i.e., thardron is analogous to water, and thardroning is analogous to raining). Therefore, people should import knowledge concerning rain and water in making a judgment concerning causal necessity and sufficiency. The observed pattern of responses should be more similar to those for the first scenario.

To summarize, it was predicted that consideration of alternative causes and disabling conditions would be greater in the familiar causal scenario than the unfamiliar one. As a result, reasoners would be less willing to accept conclusions based on the familiar scenario than on the foreign scenario. The analogy condition should fall between these two.

## Method

**Subjects.** Students ( $n = 210$ ) from introductory philosophy and cognitive psychology classes participated in the study. Participation was voluntary. None had had logic training.

**Materials.** The three scenarios presented above were typed on sheets of paper, one scenario to a sheet. Beneath the scenario were statements corresponding to modus ponens, modus tollens, denying the antecedent, and affirming the consequent. The rating scale used in the previous experiment was used here too, and it appeared beneath each argument. Examples of AC for each scenario are as follows:

*Familiar:* At dawn, Julie's roommate told her that *the streets were wet*. Can she conclude that *it rained*? Circle the response below that best reflects your decision.

*Foreign:* At dawn, Phedra's roommate told her that *the streets were sticky*. Can she conclude that *it thardroned*?

*Analogy:* At dawn, Phedra's roommate told her that *the streets were wet*. Can she conclude that *it thardroned*? Circle the response below that best reflects your decision.

The subjects were told to assume that the roommate was reliable.

**Procedure.** The subjects were run in groups prior to class lectures. They were instructed to read the story and answer the questions that followed by making a mark on the rating scale. Instructions for use of the rating scale were the same as in Experiment 1. Scenario type was treated as a between-subjects variable and argument type was treated as a within-subjects variable. There were 70 subjects in each of the three reasoning scenario conditions.

## Results

The numbers  $-3$  to  $+3$  were assigned to the rating scale as before. The ratings that subjects gave to the arguments were subjected to an ANOVA with scenario type (familiar, foreign, or analogy) as a between-subjects variable and argument type (MP, MT, DA, AC) as repeated measures. The mean ratings are illustrated in Figure 4. The main effects of causal scenario and argument type were significant [ $F(2,207) = 4.39$ ,  $MS_e = 7.10$ ,  $p < .025$ ;  $F(3,621) = 21.27$ ,  $MS_e = 2.13$ ,  $p < .001$ , respectively]. These factors did not interact with each other [ $F(6,621) = 1.31$ ,  $p > .20$ ].

Pairwise comparisons indicated that argument conclusions were given higher acceptance ratings in the foreign reasoning scenario than in the familiar (1.10 vs. .44). Ratings in the analogy reasoning scenario fell between these two and did not differ significantly from those for either the foreign or the familiar conditions (.76). As predicted, therefore, reasoners were more conservative about drawing causal inferences in the familiar case.

Pairwise comparisons also indicated that MP was given significantly higher ratings (1.45) than MT (.65), DA (.50), or AC (.46), while the latter three did not differ from each other. This effect is apparent in Figure 4; MP received consistently higher ratings than did the other arguments within each reasoning scenario, although the absolute ratings varied as a function of the scenario in which it appeared. This is consistent with the idea proposed by numerous researchers that modus ponens is part of our naive/innate reasoning apparatus (e.g., Brain & O'Brien, 1991; Rips, 1983; Smith, Langston, & Nisbett, 1992).

As in Experiment 1, the interpretations reasoners assigned to these causal conditionals was assessed by collapsing the rating scale to two points: *accept* or *don't accept/can't tell*. The patterns of responses produced by individual subjects were then tabulated. As in Experiment 1, the most preferred interpretation for all scenarios was the biconditional (accept all arguments), accounting for 37% of all observed patterns. The next most frequently observed pattern was one in which all arguments were rejected (10%). By way of contrast, the truth-functional conditional pattern accounted for only 5% of all patterns. More interesting, however, was the way in which the "accept all" and "reject all" patterns distributed themselves across reasoning scenarios. In the foreign scenario case, reasoners were 10 times more likely to accept all arguments based on the causal statement than they were to reject all arguments ( $n = 30$  vs.  $n = 3$ ). In contrast, they were a little less than twice as

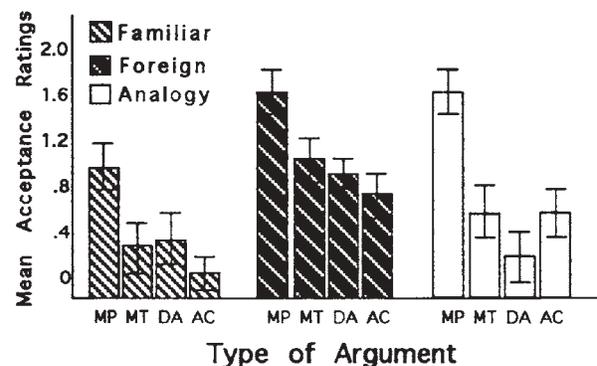


Figure 4. Mean acceptance ratings for arguments based on causal conditionals that described familiar or foreign causal relationships, or a foreign relationship that was analogous to a familiar one. The rating scale ranged from  $-3$  (very sure cannot draw this conclusion) to  $+3$  (very sure can draw this conclusion), with 0 representing *can't tell*. The arguments were: modus ponens (MP), modus tollens (MT), denying the antecedent (DA), and affirming the consequent (AC).

likely to accept all arguments than to reject them in the familiar case ( $n = 25$  vs.  $n = 13$ ). The analogy scenario produced a value between these two, with reasoners being a little more than 4.5 times more likely to accept all arguments than to reject them all ( $n = 23$  vs.  $n = 5$ ). This contingency between reasoning scenario and response pattern was significant [ $\chi^2(2) = 6.92, p < .05$ ]. Clearly, the less familiar the causal scenario, the more certain reasoners were about the causal necessity and sufficiency that characterized it.

To summarize, reasoners were more willing to draw conclusions based on unfamiliar causal scenarios than they were to draw conclusions based on familiar ones. This is consistent with the causal analysis provided here wherein causal inferences are influenced by the number of alternative causes/disabling conditions suggested by the causal scenario under consideration. If few suggest themselves, as in the foreign case, the more likely reasoners will be to accept the necessity and sufficiency of the causal relationship at face value. This produces the paradox of reasoners' being more willing to draw conclusions in situations where there is more uncertainty—that is, less information.

### GENERAL DISCUSSION

In this paper, I have provided more evidence that beliefs concerning alternative causes and disabling conditions directly impact naive causal deduction. When the reasoning context does *not* invite such considerations (foreign causal scenario) or the considerations are few in number (few alternative causes and disabling conditions), reasoners willingly draw conclusions based on the causal relationship. When the reasoning context *does* invite such considerations and the considerations are many, reasoners are far more conservative in drawing inferences. These results provide strong support for the inclusion of these factors in our naive, domain-specific causal reasoning strategies.

Furthermore, I have argued that these factors have an impact on causal deduction by influencing confidence in causal necessity and sufficiency. The presence of *disabling* conditions casts doubt on the causal *sufficiency* of the relationship in question, and the presence of *alternative causes* casts doubt on its *necessity*. In contrast, truth-functional necessity and sufficiency are based on the semantic function assigned to the conditional connective (i.e., “horseshoe”). The function states that the antecedent of a truth-functional conditional is considered sufficient for the consequent, while the consequent is considered necessary for the antecedent. Logical necessity and sufficiency are semantic relations—not pragmatic relations—between constituent propositions.

As noted by philosophers of science, however, logical sufficiency and necessity are not the same as causal sufficiency and necessity (see, e.g., Goodman, 1955; Lewis, 1973; Mackie, 1974). The former depends on the semantics of the system one constructs (e.g., first-order logic); the latter is an empirical matter that depends on

scientific theory. The unique aspect of the causal analysis offered here is the role attributed to consideration of disabling conditions and alternative causes. Disabling conditions and alternative causes have an impact on one's judgment about *causal* sufficiency and necessity, respectively, which in turn have an impact on one's deductive judgments. For example, the fact that many things, other than applying the brakes, can cause a car to slow down casts doubt on the necessity of attributing the car's slowing down to the application of the brake. It doesn't matter whether the cause is contained in the antecedent or the consequent of a conditional statement; our judgments about causal necessity will be influenced by the consideration of alternative causes, regardless of where the cause occurs in the conditional. Similarly, the fact that many things can prevent a car from slowing down even though the brake is depressed (disabling conditions) influences our judgments of causal sufficiency; depressing the brake will be considered an insufficient reason to assume that the car will slow down. Although *logical* sufficiency properly refers only to the antecedent side of a material conditional statement, *causal* sufficiency refers to likelihood of the effect's occurring in the presence of a viable cause—regardless of how the cause and effect propositions are ordered in a premise.

Various proposals have been put forth by philosophers of science for dealing with this distinction between logical and causal necessity and sufficiency. Mackie (1974) noted that the concepts of logical necessity and sufficiency cannot account for the asymmetry between causes and effects. For example, the length of a pendulum is a logically necessary and sufficient condition for the length of its arc and vice versa, yet the pendulum length is the *cause* of the arc length and not the other way around. For this reason, Mackie argued that the notion of *causal priority* must be added to analyses of causation that rely on the concepts of necessity and sufficiency. Suppes (1970) argued that our everyday conception of causality is not sharply deterministic (as are the logical concepts of necessity and sufficiency), but that it is instead probabilistic in nature. For example, accidents occur only probabilistically in the presence of reckless driving (reckless driving is neither necessary nor sufficient for an accident), yet reckless driving is nonetheless a true cause of many accidents.

This analysis suggests that the untutored reasoner's naive theories of necessity and sufficiency may exist on a continuum, as extreme points on dimensions representing beliefs about possible alternative causes and disabling conditions. If no alternative causes for an effect are possible, the cause in question is considered *necessary* for the occurrence of the effect in question, and we may confidently assert that the cause must have occurred given that the effect did. If many are possible, the cause in question is *not* considered necessary for the occurrence of the effect, and no conclusion can be drawn with certainty about the cause's role in the effect's occurrence. Similarly, if no disabling conditions are possible, the cause is considered *sufficient* to produce the effect in

question, and we may quite confidently assert that the effect will occur. In contrast, if many are possible, the cause is considered *insufficient* to produce the effect in question, and no safe conclusion can be drawn about the effect's occurrence in the presence of the cause.

This raises two possibilities. The first is that causal judgments are inherently nondeductive—that is, that even when one is reasoning with ostensibly deductive arguments, naive human reasoning is best characterized as an inductive, probabilistic process. In the ideal case, a cause is an event that is necessary and sufficient to produce a particular effect. In everyday causation, however, such perfection is rarely encountered. Instead, our naive understanding of causality may suggest the consideration of alternative causes and disabling conditions so that we can assess causal necessity and sufficiency relations between the cause and effect in question, and our decisions are made probabilistically. This interpretation is consistent with the data presented here in that acceptability ratings varied across arguments that were characterized by few to many—not one or more than one—alternatives. The second possibility, and the one championed here, is that assessment of causal necessity and sufficiency is inherently a nondeductive process, but that the outcome of this process is used by a deductive reasoning schema. Conclusions are either entailed or not entailed, and what varies is our confidence in those judgments. This interpretation is also consistent with the data presented here. More research is needed to test these competing explanations. The work presented here, however, provides unequivocal evidence that consideration of alternative causes and disabling conditions is indeed a part of our naive causal reasoning strategies.

#### REFERENCES

- BAILLARGEON, R. (1987). Object permanence in 3½- and 4½-month-old infants. *Developmental Psychology*, **23**, 655-664.
- BAUM, R. (1989). *Logic* (3rd ed.). Fort Worth, TX: Harcourt Brace Jovanovich.
- BOCCIA, M., & CAMPOS, J. J. (1989). Maternal emotional signals, social referencing, and infants' reactions to strangers. *New Directions for Child Development*, **44**, 25-49.
- BRAINE, M. D. S., & O'BRIEN, D. P. (1991). A theory of *If*: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, **98**, 182-203.
- CAREY, S. (1986). Constraints on semantic development. In W. Demopolous & A. Marras (Eds.), *Language learning and concept acquisition: Foundational issues* (pp. 154-172). Norwood, NJ: Ablex.
- CHAPMAN, L. J., & CHAPMAN, J. P. (1969). Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, **74**, 271-280.
- CHENEY, D. L., & SEYFARTH, R. M. (1985). The social and non-social world of non-human primates. In R. A. Hinde, A.-N. Perret-Clermont, & J. Stevenson-Hinde (Eds.), *Social relationships and cognitive development* (pp. 33-46). Oxford: Oxford University Press, Clarendon Press.
- CHENG, P. W., & HOLYOAK, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, **17**, 391-416.
- CHENG, P. W., & HOLYOAK, K. J. (1989). On the natural selection of reasoning theories. *Cognition*, **33**, 285-313.
- CHENG, P. W., HOLYOAK, K. J., NISBETT, R. E., & OLIVER, L. M. (1986). Pragmatic v syntactic approaches to training deductive reasoning. *Cognitive Psychology*, **18**, 293-328.
- CHENG, P. W., & NISBETT, R. E. (1991, April). *A pragmatic schema for causal deduction*. Paper presented at the meeting of the Rocky Mountain Psychological Society, Tucson, AZ.
- CHENG, P. W., & NOVICK, L. (1991). Causes versus enabling conditions. *Cognition*, **40**, 83-120.
- CHENG, P. W., & NOVICK, L. (1992). Covariation in natural causal induction. *Psychological Review*, **99**, 365-382.
- COSMIDES, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, **31**, 187-276.
- COSMIDES, L., & TOOBY, J. (1989). Evolutionary psychology and the generation culture: Part II. Case study: A computational theory of social exchange. *Ethology & Sociobiology*, **10**, 51-97.
- CUMMINS, D. D. (1995a, April). *Are pragmatic reasoning schemas innate? Some evidence*. Paper presented at the Conference on Epistemology and Evolutionary Psychology, Rutgers University.
- CUMMINS, D. D. (1995b). *Evidence of permission reasoning schemas among 3- and 4-year-old children*. Manuscript submitted for publication.
- CUMMINS, D. D., LUBART, T., ALKSNIS, O., & RIST, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, **19**, 274-282.
- DOWNING, C. J., STERNBERG, R. J., & ROSS, B. H. (1985). Multicausal inference: Evaluation of evidence in causally complex situations. *Journal of Experimental Psychology: General*, **114**, 239-263.
- EVANS, J. ST. B. T. (1989). *Bias in human reasoning*. Hillsdale, NJ: Erlbaum.
- GARCIA, J., & KOELLING, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, **4**, 123-124.
- GIGERENZER, G., & HUG, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, **43**, 127-171.
- GIROTTO, V., GILLY, M., BLAYE, A., & LIGHT, P. (1989). Children's performance in the selection task: Plausibility and familiarity. *British Journal of Psychology*, **80**, 79-95.
- GIROTTO, V., LIGHT, P., & COLBOURN, C. J. (1988). Pragmatic schemas and conditional reasoning in children. *Quarterly Journal of Experimental Psychology*, **40A**, 469-482.
- GOODMAN, N. (1955). *Fact, fiction, and forecast*. Cambridge: Cambridge University Press.
- GREENHOUSE, S. W., & GEISSER, S. (1959). On methods in the analysis of profile data. *Psychometrika*, **24**, 95-112.
- GRIGGS, R. A., & COX, J. R. (1982). The elusive thematic materials effect in Wason's selection task. *British Journal of Psychology*, **73**, 407-420.
- KELLEY, H. H. (1973). The process of causal attribution. *American Psychologist*, **28**, 107-128.
- KEPPEL, G. (1973). *Design and analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- LESLIE, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, **94**, 412-426.
- LESLIE, A. M., & KEEBLE, S. (1987). Do six-month-old infants perceive causality? *Cognition*, **25**, 265-288.
- LEWIS, D. K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.
- LIGHT, P., BLAYE, A., GILLY, M., & GIROTTO, V. (1989). Pragmatic schemas and logical reasoning in 6- to 8-year-old children. *Cognitive Development*, **4**, 49-64.
- LIGHT, P., GIROTTO, V., & LEGRENZI, P. (1990). Children's reasoning on conditional promises and permissions. *Cognitive Development*, **5**, 369-383.
- MACKIE, J. L. (1974). *The cement of the universe*. Oxford: Oxford University Press.
- MANKTELOW, K. I., & OVER, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, **39**, 85-105.
- MARKOVITS, H. (1986). Familiarity effects in conditional reasoning. *Journal of Educational Psychology*, **78**, 492-494.
- MCCLOSKEY, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.

- RIPS, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, **90**, 38-71.
- ROBERGE, J. J. (1982). Linguistic factors in conditional reasoning. *Quarterly Journal of Experimental Psychology*, **34A**, 275-284.
- RUMAIN, B., CONNELL, J., & BRAINE, M. D. S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: IF is not the biconditional. *Developmental Psychology*, **19**, 471-481.
- SEDLAK, A. J., & KURTZ, S. T. (1981). A review of children's use of causal inference principles. *Child Development*, **52**, 759-784.
- SHULTZ, T. R. (1982). Causal reasoning in the social and non-social realms. *Canadian Journal of Behavioral Science*, **14**, 308-322.
- SMETANA, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, **52**, 1333-1336.
- SMITH, E. E., LANGSTON, C., & NISBETT, R. (1992). The case for rules in reasoning. *Cognitive Science*, **16**, 1-40.
- SPELKE, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 133-169). Hillsdale, NJ: Erlbaum.
- SPELKE, E. S. (1994). Initial knowledge: Six suggestions. *Cognition*, **50**, 431-445.
- STAUDENMEYER, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults*. Hillsdale, NJ: Erlbaum.
- SUPPES, P. (1970). *A probabilistic theory of causality*. Amsterdam: Elsevier, North-Holland.
- TRONICK, E. Z., ALS, H., ADAMSON, L., WISE, S., & BRAZELTON, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, **17**, 1-13.
- TURIEL, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.
- WYNN, K. (1991). Psychological evidence against empiricist theories of mathematical knowledge. *Mind & Language*, **7**, 315-332.

## NOTES

1. The relevant statistics are as follows: For the simple interaction of argument and cause,  $F(3,33) = 6.82$ ,  $MS_e = .34$ ,  $p < .01$ . For the simple interaction of argument and disabling conditions,  $F(3,33) = 3.85$ ,  $MS_e = .41$ ,  $p < .05$ . Simple main effect analyses indicated that the number of causes influenced *only* DA [ $F(1,11) = 24.11$ ,  $MS_e = .36$ ,  $p < .001$ ] and AC [ $F(1,11) = 28.10$ ,  $MS_e = .55$ ,  $p < .001$ ]. Similarly, the number of disabling conditions influenced *only* MT [ $F(1,11) = 5.27$ ,  $MS_e = .69$ ,  $p < .05$ ]. Although the difference between the mean ratings for MP was in the predicted direction, it only achieved marginal statistical significance [ $F(1,11) = 2.76$ ,  $MS_e = .42$ ,  $.05 < p < .10$ ].

2. The relevant statistics are as follows: For the simple interaction of argument and causes,  $F(3,33) = 5.91$ ,  $MS_e = .41$ ,  $p < .001$ . For the simple interaction of argument and disabling conditions,  $F(3,33) = 5.17$ ,  $MS_e = .52$ ,  $p < .001$ . Simple main effect analyses indicated that causes influenced ratings of MP and MT *only* [ $F_s(1,11) = 17.50$  and  $14.39$ ,  $MS_e_s = .46$  and  $.62$ ,  $ps < .01$ , respectively]. Similar analyses indicated that disabling conditions influenced DA and AC [ $F_s(1,11) = 46.68$  and  $5.11$ ,  $MS_e_s = .14$  and  $.86$ ,  $ps < .001$  and  $.05$ , respectively].

**APPENDIX**  
**Conditionals on Which MP, MT, DA, and AC Arguments Were Based in Experiment 1**

| Conditonal  | Mean AC | Mean DC |
|---|---------|---------|
| Many Alternative Causes, Many Disabling Conditions                                |         |         |
| If fertilizer was put on the plants, then they grew quickly.                      | 4.2     | 3.4     |
| If the brake was depressed, then the car slowed down.                             | 4.7     | 3.1     |
| If John studied hard, then he did well on the test.                               | 3.9     | 4.4     |
| If Jenny turned on the air conditioner, then she felt cool.                       | 3.9     | 3.6     |
| Many Alternative Causes, Few Disabling Conditions                                 |         |         |
| If Alvin read without his glasses, then he got a headache.                        | 4.4     | 1.5     |
| If Mary jumped into the swimming pool, then she got wet.                          | 4.2     | 2.3     |
| If the apples were ripe, then they fell from the tree.                            | 4.2     | 1.9     |
| If water was poured on the campfire, then the fire went out.                      | 3.8     | 2.2     |
| Few Alternative Causes, Many Disabling Conditions                                 |         |         |
| If the trigger was pulled, then the gun fired.                                    | 2.0     | 2.9     |
| If the correct switch was flipped, then the porch light went on.                  | 1.8     | 2.7     |
| If the ignition key was turned, then the car started.                             | 1.6     | 3.7     |
| If the match was struck, then it lit.   | 1.9     | 3.7     |
| Few Alternative Causes, Few Disabling Conditions                                  |         |         |
| If Joe cut his finger, then it bled.  | .7      | 2.1     |
| If Larry grasped the glass with his bare hands, then his fingerprints were on it. | .8      | 1.9     |
| If the gong was struck, then it sounded.  | 2.5     | 2.0     |
| If the doorbell was pushed, then it will rang.                                    | 2.3     | 2.5     |

(Manuscript received May 4, 1994;  
revision accepted for publication October 20, 1994.)