Running head: EVERYDAY CAUSATION

Is everyday causation deterministic or probabilistic?

Caren A. Frosch[ab]

P.N. Johnson-Laird[c]

[a] University of Reading, Department of Psychology, Reading, RG6 6AL, UK.

[b] Present Address: Queen's University Belfast, School of Psychology, Belfast, UK.

c.frosch@qub.ac.uk

[c] Princeton University, Department of Psychology, Princeton, NJ 08544-1010, USA.

phil@princeton.edu

20th December 2010

Correspondence:

Caren Frosch

School of Psychology

Queen's University Belfast

Belfast

BT9 5BP

United Kingdom

Email c.frosch@qub.ac.uk

Telephone: +44 (0) 289097 6982

Fax:+44 (0)28 90975486

Abstract

One view of causation is deterministic: *A causes B* means that whenever *A* occurs, *B* occurs. An alternative view is that causation is probabilistic: the assertion means that given *A*, the probability of *B* is greater than some criterion, such as the probability of *B* given *not-A*. Evidence about the induction of causal relations cannot readily decide between these alternative accounts, and so we examined how people *refute* causal assertions. In four experiments most participants judged that a single counterexample of *A* and *not-B* refuted assertions of the form, *A causes B*. And, as a deterministic theory based on mental models predicted, participants were more likely to request multiple refutations for assertions of the form, *A enables B*. Similarly, refutations of the form *not-A* and *B* were more frequent for enabling than causal assertions. Causation in daily life seems to be a deterministic concept.

1. Introduction

The everyday concept of causation is puzzling. No-one seems sure about what it means, and some theorists even seek to outlaw it from scientific discourse (e.g., Russell, 1912-13). A more recent skeptic wrote: "There is, in fact, no such thing as cause and effect. It is a popular chimera, a vague notion that will not withstand the batterings of pure reason. It contains an inconsistent set of contradictory ideas and is of little or no value in scientific discourse" (Salsburg, 2001, p. 185-6). One reason for such claims is that common assumptions about causation *are* inconsistent (Johnson-Laird, 2006, Ch. 22). On the one hand, you assume that you can initiate a causal chain. You serve in tennis; your opponent returns the ball; and you play to and fro until the rally ends. Each serve initiates a causal chain. Nature too can intervene to initiate a causal chain: an earthquake causes a building to collapse. The notion of an intervention initiating a causal chain is plausible (see Sloman, 2005; Woodward, 2003). On the other hand, you may assume that every event has a cause (see Lewis, 1986, for the role of this assumption in reasoning about causation). You are watching TV and suddenly the screen goes blank. You infer that something has caused this event – perhaps, the set has lost power, or the system transmitting the program has gone down. Yet, if every event has a cause, then an intervention that seems to initiate a causal chain does not really do so, because it too has a cause. But, now, you are on a slippery slope back to the uncaused cause that initiated all causal chains – perhaps the one great causal chain of being, of which all other chains are mere links. Hence, either you can initiate a causal chain or else every event has a cause, but not both.

An independent question concerns, not common assumptions about causality, but the meaning of the everyday concept itself, as underlying such verbs as *push* and *pull*, and such assertions as *the moon causes the tides*. Skeptics can hardly deny the existence

of such a concept, and the question is whether it is deterministic or probabilistic. In the present paper, we outline various theoretical views about the answer, and distinguish between inductive *evidence* for causal relations and the intrinsic *meaning* of such relations. We then turn to some empirical studies of what facts individuals seek in order to refute causal assertions. We take these facts to reflect their conception of causation. Our aim is to make progress towards answering the question in the title of our paper.

The traditional view of causation is deterministic. As Hume (1748/1988, p. 115) wrote: "We may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or in other words, where, if the first object had not been, the second never had existed.*" Hume took causation to depend on nothing more than a constant conjunction of cause and effect, whereas Kant argued for a necessary connection between them, which he took to be part of the innate conception of causation. It demands "that something, *A*, should be of such a nature, that something else, *B*, should follow from it necessarily" (Kant, 1787/1934, p. 90). Mill (1874, p. 237) also held a deterministic view: "The invariable antecedent is termed the cause; the invariable consequent, the effect" (Mill, 1874, p. 237). And the reason that Russell (1912–13) argued that causation should be expurgated from philosophy was, ironically, because he *presupposed* that it was a deterministic concept, and that science demanded probabilities instead.

In the twentieth century, perhaps reflecting the irreducible probabilities of quantum mechanics, philosophers developed probabilistic accounts of causation. Reichenbach (1956) proposed such an analysis, and others followed in his steps (e.g., Salmon, 1980; Suppes, 1970, 1984). Reichenbach argued that for causation, as in *A causes B*, the following inequality should hold:

$p(B \mid A) > p(B \mid \text{not-}A)$

That is, the conditional probability of the effect *B* given the cause *A* should be greater than the conditional probability of the effect *B* given that the cause *A* did not occur. Reichenbach also noted that a cause can render irrelevant other events associated with an increase in the probability of its effect. Hence, if the probability of the effect given both the cause and the other event is the same as the probability of the effect given the cause alone, then the other event is irrelevant. Both deterministic and probabilistic views have current proponents in psychology, and we consider both sorts of theory in turn, and then why it is so difficult to decide between them from evidence about the induction of causality.

1.1. A deterministic theory of causality

The theory of mental models provides a deterministic account of the everyday meaning of causation (Goldvarg & Johnson-Laird, 2001; Johnson-Laird, 2006). *A causes B* refers to three temporally-constrained possibilities:

A	B
not-A	B
not-A	not-B

The temporal constraint is that *B* does not precede *A* in time, as corroborated in experimental studies (e.g., Bullock, Gelman, & Baillargeon, 1982; Tversky & Kahneman, 1980). When the events have occurred and are reported in the past tense, *A caused B*, the first of the possibilities above refers to a fact, and the other two cases refer to counterfactual possibilities, which support assertions such as, *if A hadn't happened then B might not have happened* (Goldvarg & Johnson-Laird, 2001; Byrne, 2005). Alternatively, when the cause failed to occur, an analogous change in reference occurs to support the counterfactual assertion, *if A had occurred then B would have occurred*.

Many theories deny that there is any distinction in meaning between causes and enabling conditions. Mill (1874) argued that the difference is capricious (see also Hart & Honoré, 1985). Cheng and Novick (1991) stipulate that the cause is inconstant and the enabling condition is constant within the relevant focal set. According to others, the cause violates a norm assumed by default whereas the enabling condition does not (see e.g., Einhorn & Hogarth, 1986; Kahneman & Miller, 1986). And, according to still another group of theorists, the cause is the factor that is relevant in any explanatory conversation: speakers describe the cause, not the enabler (Hilton & Erb, 1996; Mackie, 1980; Turnbull & Slugoski, 1988). Unlike these accounts, the model theory draws a sharp distinction between the *meaning* of causal and enabling assertions. *A enables B* refers to the following three temporally-constrained possibilities:

A	B
A	not-B
not-A	not-B

But, many assertions, such as, *a fortune enables you to live well*, have a weaker sense that is consistent with all four contingencies, i.e., even without a fortune you can live well. In daily life, however, there is often an implicature that only the antecedent, *A*, makes the consequent, *B*, possible (Goldvarg & Johnson-Laird, 2001).

To hold three distinct possibilities in mind is difficult (Bauer & Johnson-Laird, 1993; Bucciarelli & Johnson-Laird, 1999), and so a central assumption of the model theory is that individuals aim to minimize the number of models, and in particular, *mental* models normally represent explicitly only the case in which both the clauses in an assertion are true. Hence, both *A causes B* and *A enables B* have the same mental models:

A   B

. . .

where the ellipsis denotes other implicit possibilities. It follows that individuals should

not normally distinguish between the meanings of causal and enabling assertions, which

may account for the common view that they do not differ in meaning.

The model theory's concept of causation is agnostic about assumptions

concerning causation, such as whether every event has a cause or events can initiate

causal chains. However, interventions are sometimes said to have their own special logic

(e.g., Sloman & Lagnado, 2005). As an example, consider the causal assertion that eating

too much causes obesity. Granted its truth, if you were to observe that Pat isn't obese,

then you would infer that he doesn't overeat. But, suppose you learn that he takes a pill

that prevents obesity. Now, you would no longer infer from his lack of obesity that he

doesn't overeat. The pill disables the effects of overeating. No special reasoning is

needed, but just an ability to understand the premises:

Overeating causes obesity.

Taking an anti-obesity pill prevents obesity.

and to realize that the second premise takes precedence over the first (see Johnson-Laird,

2006, p. 312 et seq.).

What observations in principle refute a causal assertion? According to the model

theory, individuals grasp that a counterexample refutes an invalid inference (Johnson-

Laird & Byrne, 1991), and evidence corroborates this hypothesis (e.g., Johnson-Laird &

Hasson, 2003). A single observation of the occurrence of A without B should therefore

suffice to refute the assertion, *A causes B*. Likewise, because the theory treats *A prevents*

B as equivalent to *A causes B not to occur*, a single observation of the occurrence of A

with B suffices to refute *A prevents B*. Recently, Mandel and Vartanian (2009) have made

the same predictions about causation and prevention for similar reasons. They argue that

individuals are prone to two biases: they focus on cases in which A and B co-occur in

inferring a causal relation, which corresponds to the mental models of the concept, and they take a causal relation to mean that *A* is sufficient for *B*, and so the relation is refuted by the occurrence of *A* without *B*. It is unclear whether their account extends to enabling assertions.

Individuals often do not distinguish between causes and enabling conditions, which is why many have argued that the two concepts do not differ in meaning. The two relations have the same mental models, but cause is the stronger notion, because a claim that *A causes B*, always rules out the contingency, *A* and *not-B*, whereas the weaker interpretation of *A enables B* rules out no contingencies. The failure to distinguish between the two relations implies that if individuals are forced to think of a refutation of an enabling relation, they should tend to think of *A* and *not-B*. But, they may also realize that if *A* enables *B*, it makes *B* possible, and so what should not occur is *B* without *A*, i.e., that *not-A* and *B* refutes the enabling relation (see the three possibilities above). Likewise, a person who recognizes the distinction between causes and enabling assertions should consider all possibilities and seek an observation of *not-A* and *B* in order to refute the enabling assertion. Three predictions follow at once. First, individuals should tend to treat a single observation of *A and not-B* as refuting the assertion, *A causes B*. Second, they should treat the same observation as refuting *A enables B*, but in this case they should be more likely to seek multiple observations, because they may realize other putative refutations are possible. Third, those individuals who are able to go beyond the mental models of the enabling relation and to think of the possibilities in which the relation is false should seek a direct refutation of the form *not-A and B.* Hence, refutations of this sort should be more frequent for enabling assertions than for causal assertions.

1.2. Probabilistic theories of causation

Probabilistic theories have focused on providing an account of the induction of causal relations from frequency data. Some have explicitly endorsed a deterministic conception of causality where the induction of causation relies on probabilities, i.e. imperfect knowledge of a deterministic system. For example, Pearl (2000) presents a theory of functional causal models in which all relations are deterministic and uncertainty is introduced by unobserved and uncontrolled background variables. These background variables are not affected by the variables or the mechanisms in the causal model. He proposes that instances where the cause does not produce its effect do not necessitate a probabilistic view of causation. He argues that instead individuals accept that deliberate intervention can change the default relationship and in this way even physical laws can be modified by actions and external intervention. Consequently, *A causes B* is by default deterministic and can be refuted by a single counterexample. The causal model of an enabling condition would include another generative cause and therefore imply that the enabling condition could be present in the absence of the effect.

In contrast, other proponents of probabilistic theories view the nature of causal relations as inherently probabilistic. Suppes (1970) and Cheng and Novick (1990) treat causal assertions as meaning that the conditional probability of the effect given the cause is greater than some criterion such as the conditional probability of *effect* given *that the cause does not occur* (Cheng & Novick, 1990). The probabilities can be computed from a joint distribution, such as the following one stated with frequencies of occurrence in 100 observations for each conjunction of possibilities:

oxygen	spark	fire	20
oxygen	no spark	fire	20
oxygen	no spark	no fire	20
no oxygen	spark	no fire	20

no oxygen   no spark   no fire   20

Hence, on 20 occasions oxygen was present, there was a spark, and there was a fire.

The difference between p(fire | oxygen) and p(fire | no oxygen) – the probabilistic

contrast – is markedly positive (0.67), and so according to the probabilistic view *oxygen*

causes fire. When the contrast is markedly negative, *A prevents B*. This contrast model,

which goes back to Reichenbach (1956), fails to make the correct predictions for certain

inductions, and so Cheng (1997) proposed a "power probabilistic contrast" model (the

Power PC model) in which the contrast is normalized by dividing it by the base rate for

the effect. This factor enters into various computations in order to account for different

causal tasks (e.g., Cheng, 1997; Novick & Cheng, 2004). But, even the Power PC model

fails to account for all the experimental results (see Lober & Shanks, 2000; Perales &

Shanks, 2008). Moreover, as the example about oxygen and the fire illustrates, it can

yield implausible results. The contingencies show that, in fact, oxygen is the enabling

condition for fire – with it, there can be fire, but without it there cannot be fire, whereas a

spark is a cause of fire given the presence of oxygen – if no spark occurs in this case, fire

depends on the occurrence of an alternative cause (see Goldvarg & Johnson-Laird, 2001,

and Frosch, Johnson-Laird, & Cowley, 2007, for experimental results corroborating these

judgments).

Cheng (1997) refers to the importance of focal sets in distinguishing between

causes and enabling conditions. The causal issue under examination determines the focal

set, and so if the issue is a fire in an everyday context then the assumed background is the

normal environment in which oxygen is always present. Hence, on this account, the

probabilistic contrast depends only on the cases in which oxygen is present in the

example above. The result is a probabilistic contrast for the spark within this focal set,

but no contrast for the omnipresent oxygen. Conversely, if the question of interest is what

causes fire in an oxygen-free chamber in which sparks are produced then the focal set includes only the cases where sparks are omnipresent. However, if the causal question concerns a laboratory experiment that aims to examine what produces fire then all 100 cases are included in the focal set. In this last scenario the probabilistic analysis offers no obvious way to make the distinction between cause and enabling condition, because both the oxygen and the spark have positive probabilistic contrasts, and oxygen has the greater one. The difficulty arises because the computation of conditional probabilities loses information in the joint distribution: they alone do not suffice to reconstruct it.

Gopnik and her colleagues have argued that young children build up causal models of the world, and that some relations are probabilistic (Gopnik et al., 2004). Hence, these authors advocate the use of Bayesian nets to represent causal relations. Sloman (2005) has also advocated their use to capture adult representations of causal relations. Bayesian nets are a graphical way in which to represent variables and the conditional probabilities holding among their values. They facilitate the use of Bayes's theorem to compute posterior probabilities from the representation of prior probabilities and the relevant conditional probabilities. Pearl (2000) pioneered their use to represent the causal interrelations of events. Bayesian nets can represent deterministic relations, i.e., those with a probability of 1, but their primary use is in enabling probabilistic inferences to be drawn. They readily accommodate an 'interventionist' view of causation according to which *A causes B* means by default that a conceivable intervention that changes the value of *A*, or its probability distribution, will also change the value of *B*, or its probability distribution (Sloman, 2005; Sloman & Lagnado, 2005; Spirtes, Glymour, & Scheines, 1993; Woodward, 2003). Others have similarly argued that discoveries about structure in the world can best be understood probabilistically (e.g., Kemp & Tenenbaum, 2008). Several theorists allow for both deterministic and probabilistic causal relations

(e.g., Gopnik et al., 2004; Sloman, Barbey, & Hotaling, 2009). And Sloman et al. show how a model based on forces that work together or oppose one another (Barbey & Wolff, 2007; cf. Wolff, 2007) can also yield both deterministic and probabilistic causal relations.

Why should anyone suppose that certain causal relations are probabilistic? In our view, there are three sources for the assumption, which we examine in turn. The first is the success of quantum mechanics, which replaces determinism with irreducible probabilities. It has inculcated a probabilistic metaphysics (e.g., Suppes, 1984). But, philosophers have also adduced contrary arguments. The probabilistic account cannot explain cases in which a cause decreases the probability of an effect (see e.g., Salmon, 1980). Here is an example from Tooley (1987, p. 234–5). Disease A causes death with a probability of 0.1, and disease B causes death with a probability of 0.8. Each disease, however, confers complete immunity to the other disease. In certain circumstances, individuals are bound to contract one disease or the other. As a result, a particular person contracted disease A, and died as a result. If he hadn't caught disease A then he would have caught disease B with the probability of dying equal to 0.8. It follows that the cause of his death, disease A, in fact, reduced the probability of his dying.

A second source of the probabilistic view is the language in which many causal assertions are made, e.g.:

Following this diet causes you to lose weight.

The assertion makes a *generic* claim akin to:

Birds fly.

A generic assertion can be defined as one that contains a subject – a noun phrase or gerund – that has no explicit quantifier (see Leslie, 2008). Individuals accept generics as true, but these assertions tolerate exceptions (see Khemlani, Leslie, Glucksberg, &

Rubio-Fernandez, 2007). One of their functions is to convey generalizations that allow inferences to be made by default, e.g.:

Tweety is a bird. Birds fly. Therefore, Tweety flies.

A default inference is one that can be drawn unless there is information to the contrary, such as that Tweety is an emu. This information, however, does not refute the generic assertion, which is a useful idealization precisely because it allows default inferences. In contrast, assertions containing a universal quantifier, such as:

All birds fly

are refuted by the existence of birds that do not fly, such as emus. The same argument applies to causal assertions, such as *following this diet causes you to lose weight*. It too allows default inferences, and it tolerates exceptions, such as someone who follows the diet not losing weight. In contrast, an assertion containing a universal quantifier:

In every case, following this diet causes you to lose weight.

would be refuted by a single counterexample, and Experiment 4 below corroborates this claim. Hence, the language of causal assertions can be misleading. Generics allow causal claims to tolerate counterexamples, but it is the generic aspect of the assertion – the lack of an explicit universal quantifier, not the fact that it expresses a causal relation, that establishes this tolerance. And a tolerance for counterexamples suggests a probabilistic interpretation.

A third source of the probabilistic view is methodological. Systematic evidence pertinent to causation is often statistical, in part because noise and erroneous observations are bound to occur, and in part because hidden causes and disabling factors may be uncontrolled in samples of data. For instance, if you observe that 99 out of 100 smokers develop cancer, whereas only 9 out of 100 non-smokers from the same population do, then you have prima facie evidence that smoking causes cancer. The probabilistic view

implies that such data are the end of the story. But, the deterministic view implies that individuals should think that one or more hidden disabling components are at work sparing some of those who smoke from developing cancer, and one or more hidden alternative causes leads to cancer amongst those who do not smoke. Experiments show that individuals induce causal relations even when there are exceptions, that is, the covariation of cause and effect is less than perfect (e.g., McArthur, 1972; Shanks, 2004). And they also show that individuals attribute exceptions to disabling conditions and alternative causes (see Cummins, 1995, 1998; Luhmann & Ahn, 2003, 2005), and that even children distinguish between the two (Schulz & Sommerville, 2006).

A factor that strengthens an inductive inference of a cause is the existence of an obvious mechanism, and some causal theories place this factor at the center of induction (e.g., Ahn & Bailenson, 1996; Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Sloman, 2005). Yet, mechanisms cannot be a necessary component for inferring a causal relation, because individuals readily infer causes when they have no knowledge of any underlying mechanism, as in the epidemiological studies of smoking and cancer. This point goes back to Hume (1739/1978). As he anticipated, no-one – as far as we know – has formulated a satisfactory definition of a causal power, force, or mechanism, that makes no reference, not even an implicit one, to causality. And, as he also anticipated, people treat chance as a case of a hidden cause. Some of those who have studied hidden or alternative causes defend a deterministic view of causation (e.g., Schulz & Sommerville, 2006), whereas others defend a probabilistic view (e.g., Gopnik et al., 2004).

The induction of a causal claim, whether individuals regard causation as deterministic or probabilistic, depends on evidence, and evidence almost always contains "noise" of one sort or another. Hence, no inconsistency occurs in inducing a

deterministic causal relation between two events even though the data contains occasional counterexamples. We are therefore unable to devise any simple inductive experiment to address the contrast between the two conceptions of causation. Evidence for causation is usually probabilistic. However, if the resulting *concept* of causation is a matter of a conditional probability larger than some criterion, it should sometimes be ascribed to correlated events when the conditional probability is less than one. And it follows that the mere observation of a single counterexample does not refute the causal relation. Refutation also calls for probabilistic evidence: if individuals are asked in general terms for the refutation of causal claims, they should usually demand evidence of frequent counterexamples, and reject a single counterexample as sufficient to refute such claims. This asymmetry between induction and refutation led us to examine the contrast between the two conceptions of causation by examining our participants' judgments about what, in principle, refutes causal claims. It seems that the results of inductive studies have no uncontroversial implications for the question that concerns us. We therefore turned our attention to the refutation of causal assertions.

1.3. The refutation of causal claims

A clear implication of the probabilistic view is that a causal assertion, *A causes B*, should tolerate exceptions without the need for explanation. The assertion means that there is a high conditional probability of *B* given *A*, and so the observation of a single case of *A* without *B* does not refute the assertion. Its refutation calls instead for a series of observations that yield a reliable probability that is robustly smaller than the assumed value of p(B | A), and perhaps not significantly different from the probability of p(B | not-A). Likewise, the refutation of the assertion *A enables B* should call for a series of observations that are robustly smaller from the assumed value of a given probability,

though in this case quite what the conditional probability should concern, and whether it should differ from p(B | A), is unclear.

Other theories lack unequivocal predictions about the refutations of causal assertions. Thus, Wolff (personal communication 06/12/2008) acknowledges that the force dynamics theory makes no clear predictions about how causes and enablers should be refuted, but suggests that a combination of the theory and lexical semantics might predict that an assertion of the form, *A causes B*, is refuted by an observation of *A* and *not-B*, whereas an observation of *not-A* and *B* would be more damaging for an enabling relation. The theory accordingly makes much the same predictions as the model theory. But, other current theories take a different point of view. Sloman's (2005) causal model theory stresses the importance of mechanisms, and so it implies that refutations should establish either the absence of a mechanism relating cause to effect or that the mechanism is malfunctioning. Hence, causes are refuted by the absence of an enabler or the presence of a disabler; and enablers are refuted by the absence of additional enablers. In some cases, a single refutation suffices; it depends on how many other causes, enablers, or disablers, an assertion brings to mind (Sloman, personal communications, 06/24/2008 and 09/29/2008). Sloman et al (2009) have recently proposed that an enabler implies an accessory cause, i.e., the causal role of an enabler can only be fully understood with reference to the cause it relates to. Additionally, a particular enabler is regarded as an instance of a category of events that come together with a cause to produce a particular outcome. Hence, we may expect more multiple refutations for enablers than for causes and in particular these refutations would refer to the missing accessory cause.

In contrast, as we showed earlier, the model theory makes three immediate predictions about the refutations of causal assertions: 1. a single observation of *A and not-B* should refute the assertion, *A causes B*; 2. it should refute *A enables B*, but this

relation should be more likely to call for multiple refutations; and, in particular, 3. refutations of the form *not-A and B* should be more frequent for enabling assertions than for causal assertions. We carried out four experiments to test these predictions.

2. Experiment 1

In our initial study, the participants' task was to respond spontaneously to an open-ended question asking them what would refute assertions concerning cause, prevention, or an enabling condition. The assertions were biased against the model theory's predictions because they were generics (in the sense defined in 1.2, i.e., their subjects had no explicit quantifiers), e.g.: *Regular exercise of this sort causes a person to build muscle*. Such an assertion tolerates exceptions precisely because it lacks a quantifier.

2.1. Method

2.1.1. Design

The participants acted as their own controls. They were asked for refutations of 15 assertions, five each for *causes*, *enables*, and *prevents*. Each of the contents occurred in all three sorts of assertion, but every participant saw just one assertion containing a given content. The assertions were presented in a different random order to each of the participants.

2.1.2. Participants

A sample of 18 Princeton University students and staff (mean age = 22 years) from a variety of disciplines received $10 to participate in the experiment.

2.1.3. Materials

The materials were 15 assertions – three each from five domains: physiological, physical, mechanical, socio-economic, and psychological. An assertion concerning physiology, for example, was: *Regular exercise of this sort causes a person to build muscle*. The full set of 15 assertions is in the Appendix (the first three items in each domain). Wolff and Song (2003) showed that their participants distinguished between causing and enabling relations whether the causal relations were expressed using *causes, forces*, or *makes*, and whether enabling relations were expressed using *enables, allows*, or *helps*. There are subtle differences in meaning amongst these verbs, and so we used the most general verbs: *causes*, *enables*, and *prevents*. Individuals are more likely to generate counterexamples, single or multiple, when they refute another person's claim (Cowley & Byrne, 2005; submitted; Sanbonmatsu, Posavac, Vanous, & Ho, 2005), and so each assertion was presented as made by a different person.

2.1.4. Procedure

The experiment was conducted online at Princeton University. Participants received the following instructions:

> You will be presented with fifteen statements made by various people and for each of them you will be asked to consider whether there could be evidence that demonstrates that the person making the statement is wrong. Your task will consist of typing your responses for each of the statements.

Each statement was presented on a separate screen together with a request to describe the kind of evidence that would refute this statement, e.g.:

> Julie says:

> Regular exercise of this sort enables a person to build muscle.

What sort of evidence would you require to refute this assertion? Please describe one or more distinct possibilities that would show that Julie's assertion is false.

2.2. Results and discussion

We classified the participants' responses according to the number of refutations (single or multiple) that they required and the type of refutations that they required. A second judge who was blind to the hypotheses also classified the responses, and the 9% of disagreements were resolved through discussion. The classification for the number of refutations was stringent: only refutations that explicitly referred to one observation were classified as 'single', e.g.:

Someone who engages in this sort of regular exercise who is unable to build muscle as expected.

Assertions such as:

I would have people perform the exercise and see whether they consistently built muscle or not

were classified as calling for multiple refutations. The majority of responses for each of the three sorts of assertions called for single refutations: 70% for *causes*, 77% for *enables*, and for 79% for *prevents* (Wilcoxon tests, z = 2.43, p < .01; z = 2.52, *p < .01*; *z* = 3.37, *p* < .005, respectively, one tail tests). There was no reliable difference in the number of single refutations requested for *causes* and *enables* (Wilcoxon test, *z* = .78, *p* > .22).

Figure 1 presents the percentages of the main sorts of refutation: *A and not-B*, *not-A and B*, and *A and B*, for the three sorts of relations. The figure omits the 12% of miscellaneous responses, such as requests for additional information, which fell outside the three main categories. As Figure 1 shows, the principal refutations of *A causes B* and

A enables B were *A* and *not-B,* and the principal refutations of *A prevents B* were *A* and *B*. Although *enables* was slightly more likely to elicit refutations of *not-A* and *B*, the difference was not reliable. In sum, the experiment showed the predicted preference for single refutations over multiple refutations in an open-ended task. It yielded the expected difference between *causes* and *prevents*, but it failed to establish the predicted differences between *causes* and *enables*. However, with an open-ended question, the participants may have been tempted to provide just a single representative piece of evidence that would refute the claims. Because *causes* and *enables* have the same mental models, this evidence would be the same for both sorts of relation.

[FIGURE 1 ABOUT HERE]

3. Experiment 2

One reason for the preference for single refutations in the previous experiment may have been the open-ended nature of the task. Hence, in the present experiment, the participants were asked instead to make an explicit judgment about whether or not one observation would suffice to show that a claim was wrong. And, because the previous study clearly distinguished between *prevents* and the other two verbs, the present experiment focused on the potential difference between *causes* and *enables*. It provided additional paraphrase of the meaning of each verb, i.e., "a cause brings about an outcome," whereas "an enabler makes the outcome possible". Nearly all the refutations of these two relations in the previous experiment were either *A and not-B* or *not-A and B*. Hence, the response format in the present experiment was a forced choice between these two options followed by a question about whether this observation was sufficient to refute the relation.

3.1. Method

3.1.1. Design

The participants acted as their own controls and responded to sixteen assertions that described eight causes and eight enablers. Each of the contents, which were the same as those in the previous experiment, occurred in both sorts of assertion, but every participant saw just one assertion containing a given content. The assertions were presented in a different random order to each of the participants.

3.1.2. Participants

A new sample of 20 Princeton University students and staff (mean age = 23 years) took part in the experiment and received $10 for their participation.

3.1.3. Procedure and materials

The experiment was conducted online at Princeton University. The participants received the following instructions:

> You will be presented with sixteen statements made by various people and for each of them you will be asked to consider which piece of evidence would demonstrate best that the person making the statement is wrong.

The materials were based on those used in Experiment 1 with the following changes. The assertions concerned only *causes* and *enables*. They contained a final clause paraphrasing their meaning, e.g.:

> Regular exercise of this sort causes a person to build muscle, i.e. brings about muscle growth.
>
> Regular exercise of this sort enables a person to build muscle, i.e. makes muscle growth possible.

Each assertion was followed by the question:

> Which of these two possibilities provides more decisive evidence against this
> assertion?'

and two options of the form, *A and not-B*, and *not-A and B*, e.g.:

> A person engaged in regular exercise of this sort and yet did not build muscle.

> A person did not engage in regular exercise of this sort and yet built muscle.

Finally, after they had made the choice between the options, they answered the question:

> Would this observation suffice to show that the claim is false? If not, what other
> observation is necessary?

3.2. Results and discussion

Overall, the results corroborated each of the model theory's predictions. The majority of refutations were *A and not-B* (82.5%). Every single participant chose *A and not-B* more often than *not-A and B* (Binomial $p = .05^{20}$), but they selected an observation of *not-A and B* more often for *enables* (25%) than for causes (10%, Wilcoxon, $z = 2.53$, $p < .01$). On 87% of the trials, the participants responded to the second question that one observation was sufficient to refute the assertion, and every single participant did so more often than not (Binomial $p = .05^{20}$). But, as the model theory also predicts, they were more likely to require multiple observations for *enables* (21% of trials) than for *causes* (6% of trials; Wilcoxon test, $z = 2.35$ $p < .01$). In total, the participants requested multiple refutations on only 43 trials out of a total of 320 trials (10 for *causes* and 33 for *enables*). On these 43 trials, they requested further observations of the same sort on 10 trials, of the opposite sort on 10 trials (i.e. *A and not-B* if they had originally selected *not-A and B* and vice versa), and of other

factors on 14 trials. On the remaining 9 trials, they stated that it was impossible to refute the assertion, restated the causal claim, or made unclassifiable assertions.

4. Experiment 3

Although the previous experiment corroborated the model theory's predictions, we suspected that the formulation of the problems still left room for improvement. Hitherto, we had rotated the contents over both the causing and enabling assertions, and so we used materials that were feasible in both sorts of assertion. Hence, a stronger test of the predicted difference between causes and enablers is to select materials that are typically interpreted as causal or else as enabling and so we included four additional items: a weak causal assertion, which allows for alternative causes; a strong causal assertion, which does not; a strong enabler, which is necessary for the effect; and a weak enabler, which is consistent with all four contingencies (see the Introduction). A single observation of *A and not-B* refutes both sorts of causal assertion. A single observation of *not-A and B* refutes a strong enabler, but does not refute a weak enabler. In order to rule out any potential artefact from our previous use of online experiments, the present experiment was carried out with individual participants face to face. Similarly, the instructions provided examples of the contrast between causal and enabling relations.

4.1. Method

4.1.1. Design**,** Materials**,** and Procedure

The design was the same as in Experiment 2. The materials were similar except for the use of the four additional assertions at the end of the experiment: a weak cause, a strong cause, a strong enabler, and a weak enabler (see the Appendix). The materials were

presented in booklets, which consisted of an instruction sheet and the 20 assertions. In addition to the previous instructions, we added the following explanation:

> The statements you will be shown contain the terms 'cause' and 'enable'. Here are some examples of events that cause an outcome compared to events that enable an outcome:

> Jumping into a swimming pool CAUSES a person to get wet.

> The presence of oxygen ENABLES a fire to occur.

> As this example shows, unlike a cause, an enabler merely makes it possible for an event to occur.

4.1.2. Participants

Twenty-one University of Reading undergraduate students (mean age = 22 years) received £5 for participating in the experiment.

4.2. Results and discussion

As in the previous experiment, every single participant chose a single observation to refute the assertions more often than not (Binomial $p = .05^{20}$). But, they were more likely to do so for causal assertions (91% of trials) than for enabling assertions (63% of trials, Wilcoxon test, $z = 3.35$, $p = .0005$). As the model theory also predicts, they chose A and $not\text{-}B$ predominantly to refute both sorts of relation (18 out of the 21 participants did so, and the remainder were ties, Binomial $p = .05^{18}$). Once again, they chose $not\text{-}A$ and B more often to refute enabling relations (38% of trials) than to refute causal relations (12% of trials; Wilcoxon test, $z = 2.5$, $p < .01$). In order to refute a probabilistic relation, it is

appropriate to seek multiple observations of the same sort. For causes, however, none of the participants wanted observations of the same sort, on 2% of trials they requested information about the other sort of observation, i.e., they switched from *A and not B* to *not-A and B*; and on 7% of trials they wanted entirely different sorts of information. Their requests were different for *enables*, though the numbers were too small for statistical analysis. But, as a weak interpretation requires, the participants wanted more information of the same sort on 6% of trials, information about the other sort of observation on 6% of trials, and on 25% of trials they wanted entirely different sorts of information.

The four additional assertions with appropriate contents for *causes* and for *enables* also yielded the predicted results: the participants most frequently chose *A* and *not-B* to refute both causal assertions (90% for the strong cause and 95% for the weak cause), but they were more likely to select *not-A* and *B* for both enabling assertions (57% of the strong enabler and 38% for the weak enabler). And once again, the majority of responses were for single refutations, except, as was to be expected, for weak enablers, which are compatible with all four contingencies.

5. Experiment 4

A deterministic view of causation implies that a generic causal assertion, such as: *contact between these two sorts of substance causes an explosion to occur*, should be treated similarly to a universal causal assertion: *in every case, contact between these two sorts of substance causes an explosion to occur*. Individuals should treat both as refuted by a single counterexample. In contrast, a merely existential assertion: *in some cases, contact between these two sorts of substance causes an explosion to occur,* invites a probabilistic interpretation and individuals should require multiple counterexamples in order to refute it. If causation is itself a probabilistic concept then no reason exists – as far as we can

discern – for any difference in performance over the three sorts of assertion. Even the universal assertion should be equivalent to the following paraphrase: *in every case, there is a high conditional probability of an explosion given contact between these two sorts of substance*. Its refutation calls for a series of trials in which contact is seldom, if ever, followed by an explosion. At the very least, the probabilistic account implies that this sort of interpretation should occur for generic assertions. When an inconsistency occurs between the implications of a conditional claim and the facts of the matter, individuals seek an explanation that resolves the inconsistency (see, e.g., Johnson-Laird, Girotto, & Legrenzi, 2004), and they even rate such an explanation as more probable than a minimal revision to the premises (Khemlani & Johnson-Laird, 2010). Hence, individuals should follow up counterexamples to universal or generic claims with a search for an explanation for the inconsistency. But, no such search should occur for existential assertions, which obviously tolerate counterexamples. The experiment tested these predictions using the three sorts of assertions.

5.1. Method

5.1.1. Design and materials

The participants acted as their own controls, and were presented with 10 universal assertions containing the quantified phrase, *in every case*, 10 generic assertions with no quantified phrase, and 10 existential assertions containing the quantified phrase, *in some cases*. Here are examples of the same content in each of the three sorts of assertion:

In every case, regular exercise of this sort causes a person to build muscle, i.e. brings about muscle growth.

Regular exercise of this sort causes a person to build muscle, i.e. brings about muscle growth.

In some cases, regular exercise of this sort causes a person to build muscle, i.e. brings about muscle growth.

The three sorts of assertion were in blocks presented in all six counterbalanced orders, and the assertions within each block were presented in a different random order to each participant. Half of the assertions of each sort used the verb *causes,* and the other half of the assertions used the verb *enables*. We devised a further 14 contents that were added to those used in the previous experiments, which resulted in six items in each of the five domains (see the Appendix for a full list of the contents). Each participant received all 30 items across the three conditions, and we ensured that three items from each domain described a causal relation and three items described an enabling relation. The 30 contents were rotated over the problems so that in the experiment as a whole each content occurred equally often in the six different sorts of assertion (three sorts of quantifier x two sorts of relation).

5.1.2. Participants

Thirty-four University of Reading staff and students (mean age = 29 years) participated in the experiment in return for £8.

5.1.3. Procedure

Participants were tested individually and were given the same instructions as in the previous experiment. The first question about the sort of observation was identical to the previous two experiments, but we refined the question about the number of observations to draw attention to the fact that the participants had so far considered only a *single* observation, and we asked them in more detail about what other observation would be necessary. An example of a trial is as follows:

Following this diet causes a person with this sort of metabolism to lose weight, i.e. brings about weight loss.

Which of these two possibilities provides more decisive evidence against this statement?

☐ A person with this sort of metabolism followed this diet and yet did not lose weight.

☐ A person with this sort of metabolism did not follow this diet and yet lost weight.

Would this single observation suffice to show that the claim is false?

☐ Yes

☐ No

If not, what other observation(s) is necessary? Please specify whether you would want to see more of the same (how many?) or a different kind of observation (what sort?)

5.2. Results

Figure 2 presents the percentages of selections of single and multiple refutations for each of the six conditions in the experiment. Overall, the participants tended to choose more single refutations for generic and universal assertions (60%) than for existential assertions (28%; Wilcoxon test $z = 4.64$, $p < .001$). Likewise, they tended to choose more single refutations for causal relations (58%) than enabling relations (40%; Wilcoxon test $z = 3.2$, $p < .0025$). But, as the Figure shows, there was a significant interaction, the difference in the number of single refutations between causal and enabling relations was clear for universal and generic assertions, but non-existent for existential assertions (Wilcoxon test $z = 3.92$, $p < .001$) The experiment accordingly corroborated the crucial prediction: generic assertions tend to be treated as akin to universally quantified assertions rather than as akin to existentially quantified assertions.

[FIGURE 2 ABOUT HERE]

[FIGURE 3 ABOUT HERE]

Figure 3 presents the percentages of *A and not-B* and *not-A and B* refutations for each of the six conditions in the experiment. In every condition, most refutations were of the form *A and not-B*, and the participants tended to choose more *A and not-B* refutations for generic and universal assertions (82%) than for existential assertions (77%; and the difference was marginally significant, Wilcoxon test $z = 1.57$, $p = .058$). Likewise, they tended to choose more *A and not-B* refutations for causal relations (87%) than for enabling assertions (74%; Wilcoxon test $z = 3.12$, $p < .0025$). But, as the Figure shows, there was again a significant interaction, the difference in the number of *A and not-B* refutations between causal and enabling assertions was clear for universal and generic assertions, but almost non-existent for existential assertions (Wilcoxon test $z = 2.53$, $p < .015$). Finally, when the participants requested multiple refutations, they tended to seek other information, such as disabling conditions, for universal and generic assertions (59%), but not for the existential assertions (38%, Wilcoxon test $z = 2.94$, $p < .005$). In the latter case, as befits a probabilistic interpretation, they tended to seek more observations of the same sort (51%). If sceptics worry that our experimental procedure or contents were somehow slanted towards deterministic interpretations, the present experiment should convince them. The participants selected many multiple refutations for existential assertions for causes (see Figure 2), but for generics and universal assertions about causes they showed a clear preference for single refutations (74% of trials).

6. General Discussion

Many views about causation exist in cognitive science, but one central question about the everyday concept is embodied in the title of the present paper: is the concept deterministic or probabilistic? Previous experimental studies have provided evidence that some authors take to imply a deterministic concept, whereas others take it to imply a probabilistic concept. And, as we argued earlier, studies of the induction of causal relations have not resolved the controversy, if only because it is difficult to distinguish between the nature of evidence in an induction and the nature of the resulting concept. We accordingly examined what individuals take to *refute* a causal assertion, such as:

Pushing the lid causes this type of container to open

and to refute an enabling assertion, such as:

Pushing the lid enables this type of container to open.

Refutations reveal the conditions under which assertions are false, and accordingly reveal the nature of their underlying concepts.

A deterministic theory based on mental models, which we described earlier, predicts that individuals should seek a single observation of *A and not-B* in order to refute the assertion, *A causes B*. They should also seek it to refute *A enables B*, because the assertion has the same mental models as the causal assertion, but they should be more likely to seek multiple observations, because they may realize other putative refutations are possible. And those individuals who are able to consider all the possibilities to which an enabling assertion refers should seek a direct refutation of the form *not-A and B*. The theory also postulates a temporal order: B cannot precede A in time, but leaves open several metaphysical issues, such as whether all events have a cause.

A probabilistic view of causation, which we also reviewed earlier, makes quite different predictions (e.g., Reichenbach, 1956; Suppes, 1970; Cheng, 1997). It treats *A*

causes B as equivalent to an assertion that the conditional probability of *B* given *A* is higher than some criterion, with the additional requirements of causes preceding their effects and conditional independence or an assumption of causal mechanism. Hence, the way to refute the assertion is to show that this conditional probability is not high, and such a refutation calls for a series of observations in which *B* fails to occur in many, or most, cases in which *A* occurs. Of course, some proponents of probabilistic theories could argue that if the probabilities were set to 1 then they could account for the finding that the participants preferred single refutations to multiple refutations. But, such a rebuttal still begs the question as to why we see such an overwhelming preference for single refutations across a broad set of domains from which the causal statements were drawn. If the everyday conception of causality is probabilistic then, at the very least, we would expect to see a greater mixture of responses to the different statements. The one probabilistic account that is consistent with our results is due to Pearl (2000), because he postulates that functional causal models in which all the relations are deterministic.

It is difficult to know what probabilistic theories predict about the refutation of enabling assertions, *A enables B*, because they have had so little to say about them. One early view, however, is that an enabler is constant in the focal set of events, whereas the cause is inconstant (Cheng & Novick, 1991). This view would imply that a single case in which the enabler did not occur in the relevant situation would refute an enabling assertion, i.e., the opposite prediction to the deterministic theory based on mental models. However, enablers do not always appear to be invariable. Consider this scenario, for instance:

> Mary threw a lighted cigarette into a bush. Just as the cigarette was going out, Laura deliberately threw petrol on it. The resulting fire burnt down her neighbor's house.

Participants in an experiment identified Mary's action of throwing the cigarette into the bush as enabling the fire to occur, and Laura's action of throwing petrol onto the cigarette as its cause (Frosch, et al., 2007). Yet, the lighted cigarette is not constant, it was in the process of going out just as Laura doused it with petrol. So, the safest course is to leave open the question of what, if anything, Cheng's probabilistic theory implies about the refutation of enabling assertions. Spellman (1997) distinguishes between causes and enabling conditions by suggesting that a cause increases the probability of the outcome to a greater degree than the enabler does. However, in a different, but related question where participants were asked to assess the likelihood that removing the cause or the enabling condition would change the outcome Frosch, Egan and Hancock (in revision) found no reliable difference in ratings for the two sorts of event.

When individuals responded freely to open-ended questions, they tended to call for single refutations of causal, enabling, and preventative assertions (Experiment 1). One reason could have been that they made only minimal response to such questions. When the assertions contained a paraphrase of the meaning of *cause* as "bringing about" an event, and of *enable* as "making possible" an event, participants discriminated between the two. And, as the model theory predicts, they tended to require single refutations for causal assertions but an increased proportion of multiple refutations for enabling assertions. Similarly, they were more likely to require observations of *not-A and B* to refute enablers than to refute causes. This pattern of results was robust, occurring both in a Web study at Princeton University (Experiment 2), and in a face-to-face experiment at Reading University, UK (Experiment 3). In these studies, the assertions were couched as "generics" (see Leslie, 2008), i.e., they were assertions without explicit quantifiers. Generics, whether they concern causation or quite different matters, tolerate exceptions. Hence, people readily assent to the assertion: *ducks lay eggs*, even though they know that

only female ducks do (Khemlani et al., 2007). A generic concerning causation, such as our example: *pushing the lid causes this type of container to open*, also tolerates exceptions. Hence, our use of generics in the previous experiments might have predisposed the participants to probabilistic interpretations. We therefore compared generics with universally quantified assertions, such as:

> In every case, pushing the lid causes this type of container to open.

We also compared them with existentially quantified assertions, such as:

> In some cases, pushing the lid causes this type of container to open.

The model theory predicts that generics should elicit the same patterns of judgments as universals, whereas existential assertions should elicit refutations akin to explicitly probabilistic assertions. But, if causation is an intrinsically probabilistic concept, then no basis exists for distinguishing amongst the three sorts of assertions. Even a universal assertion should be equivalent to the following assertion:

> In every case, this type of container has a high probability of opening given that
> the lid is pushed.

The results corroborated the deterministic account (Experiment 4), even though the participants were explicitly asked – as in Experiments 2 and 3 – whether they would want to observe more outcomes of the same sort, and whether they needed any other sort of evidence.

What is wrong with the hypothesis that the everyday concept of causation is deterministic? In our view, two main arguments can be made against it. The first argument is that individuals make causal assertions even when they know that there are exceptions, e.g., *smoking causes lung cancer*. A rebuttal is that the assertion about smoking is a generic, and whether we couched causal assertions as generics or as universal claims, such as: *in every case, smoking causes lung cancer,* our participants

required only a single counterexample – a smoker who did not contract cancer – to refute the assertion. The second argument is that individuals induce causal relations even when they know that the correlation between cause and effect is less than perfect. The rebuttal is that these inductions merely reflect the fact that individuals readily envisage both hidden causes that can disable a principal cause and alternative causes that can bring about the effect in its absence (for a defence of this view, see, e.g., Luhmann & Ahn, 2005). We know of no other case against a deterministic concept of causation.

What is wrong with the hypothesis that the everyday concept of causation is probabilistic? In our view, three main phenomena count against it. The first phenomenon is that when probabilities are held constant, a manipulation of content can yield different attributions of causality (Legrenzi & Sonino, 1994; White, 1995; Koslowski, 1996). Such results are inexplicable if causation is equivalent solely to an assertion about a conditional probability. The second phenomenon is that participants tend to draw definite conclusions from causal premises, such as:

Eating protein will cause her to gain weight.

She eats protein.

Will she gain weight?

The participants had three response options:

Yes. No. Perhaps yes, perhaps no.

And the majority responded, "yes" (Goldvarg and Johnson-Laird, 2001, Experiment 3). But, if a causal assertion were merely an assertion of a high conditional probability, then they should have responded, "Perhaps yes, perhaps no".

The third phenomenon is the one that we have reported here. In four experiments using different procedures and a variety of contents, the participants overwhelmingly responded that a single case of *A and not-B* refuted a causal assertion, *A causes B*. If the

probabilistic view were correct, then they should instead have required multiple observations of this counterexample. Our evidence about enabling assertions is less decisive. It corroborates the model theory, but it isn't clear what the probabilistic view predicts about such assertions. In sum, when individuals make inferences from causal assertions, their inferences are not probabilistic. When they think about what refutes causal assertions, their responses reflect a deterministic concept. We conclude that the great philosophers – Hume, Kant, Mill, and Russell – may have been right to defend this concept of everyday causation.

Acknowledgments

References

Ahn, W., & Bailenson, J. (1996). Causal attribution as a search for underlying mechanism: an explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology*, 31, 82–123.

Barbey, A., & Wolff, P. (2007). Learning causal structure from reasoning. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 713-718). Austin, TX: Cognitive Science Society.

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science, 4*, 372-378.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23*, 247-303.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In Friedman, W.J. (Ed.) *The developmental psychology of time*. Orlando, FL: Academic Press. Pp. 209-254.

Byrne, R.M.J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT.

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

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58*, 545-567.

Cheng, P.W., & Novick, L.R. (1991). Causes versus enabling conditions. *Cognition, 40*, 83-120.

Cowley, M.*, & Byrne, R. M. J. Hypothesis falsification and opponents in the 2-4-6 task, under review.

Cowley, M. & Byrne, R. M. J. (2005). When falsification is the only path to truth. In B. G. Bara, L. Barsalou, & M. Bucciarelli (eds.). *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 512-517). Mahwah, NJ: Erlbaum.

Cummins, D. D. (1995). Naive theories and causal deduction. *Memory & Cognition*, 23, 646–658.

Cummins, D. D. (1998). The pragmatics of causal inference. *Proceedings of the twentieth annual conference of the cognitive science society* (p. 9).

Einhorn, H.J., & Hogarth, R.M. (1986) Judging probable cause. *Psychological Bulletin*, 99, 3-19.

Frosch, C.A., Egan, S.M., & Hancock, E.N. (In revision). Why counterfactual thoughts focus on enablers.

Frosch, C.A., Johnson-Laird, P.N., & Cowley, M. (2007). It's not my fault, Your Honor, I'm only the enabler. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (p. 1755). Austin, TX: Cognitive Science Society.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naive Causality: a mental model theory of causal meaning and reasoning. *Cognitive Science*, *25*, 565-610.

Gopnik, A., Glymour, C., Sobel, D.M., Schulz, D.E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 1-31.

Hart, H.L.A., & Honoré, A.M. (1985). *Causation in the law*. Second Edition. Oxford: Clarendon Press. (First edition published in 1959.)

Hilton, D.J., & Erb, H-P. (1996). Mental models and causal explanation: Judgements of probable cause and explanatory relevance. *Thinking & Reasoning*, 2, 273-308.

Hume, D. (1978). *A treatise on human nature*. Selby-Bigge, L.A. (2nd ed.). Second ed. Oxford: Oxford University Press. (Originally published 1739.)

Hume, D. (1988). *An enquiry concerning human understanding.* Ed. A. Flew. La Salle, IL: Open Court. (Originally published 1748.)

Johnson-Laird, P. N. (2006). *How we reason*. Oxford: Oxford University Press.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991) *Deduction*. Hillsdale, NJ: Erlbaum.

Johnson-Laird, P.N., & Hasson, U. (2003) Counterexamples in sentential reasoning. *Memory & Cognition*, 31, 1105-1113.

Johnson-Laird, P.N., Girotto, V., & Legrenzi, P. (2004) Reasoning from inconsistency to consistency. *Psychological Review*, 111, 640-661.

Kahneman, D., & Miller, D.T. (1986). Norm theory: Comparing reality to its alternative. *Psychological Review*, 93, 75-88.

Kant, I. (1934). *The critique of pure reason*. Second Ed. Trans. Meiklejohn, J.M.D. London: Dent, 1934. (Originally published 1787.)

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, 105, 10687-10692.

Khemlani, S. & Johnson-Laird, P. N. (2010). The need to explain. *in submission.*

Khemlani, S., Leslie, S. J., Glucksberg, S., & Rubio-Fernandez, P. (2007). Do ducks lay eggs? How people interpret generic assertions. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 395-401). Austin, TX: Cognitive Science Society.

Koslowski, B. (1996). *Theory and evidence: the development of scientific reasoning*. Cambridge, MA: MIT Press.

Lagnado, D. A., Waldmann, M. R., Hagmayer Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal*

learning: Psychology, philosophy, and computation. Oxford: Oxford University Press, pp. 154-172.

Legrenzi, P., & Sonino, M. (1994). Psychologistic aspects of Suppes's definition of causality. In Humphreys, P. (Ed.) *Patrick Suppes: scientific philosopher*, Vol. 1 (pp. 381–399). The Netherlands: Kluwer.

Leslie, S. J. (2008). Generics: cognition and acquisition. *Philosophical Review*, *117*, 1-47.

Lewis, C. (1986). A model of mental model construction. *Proceedings of CHI '86 Conference on Human Factors in Computer Systems*. New York: Association for Computing Machinery. 306-313.

Lober, K. & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, 107, 195-212.

Luhman, C. C., & Ahn, W. K. (2003). Evaluation the causal role of unobserved variables. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 734-739). Mahwah, NJ: Erlbaum.

Luhman, C. C., & Ahn, W. K. (2005). The meaning and computation of causal power: Comment on Cheng (1997) and Novick and Cheng (2004). *Psychological Review*, *112*(3), 685-693.

Mackie, J.L. (1980). *The cement of the universe: A study in causation*. Second edition. Oxford: Oxford University Press.

Mandel, D.R., & Vartanian, O. (2009). Weighting of contingency information in causal judgement: Evidence of hypothesis dependence and use of a positive-test strategy. *Quarterly Journal of Experimental Psychology*, 62, 2388-2408.

McArthur, L. (1972). The how and what of why: some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 22, 171-193.

Mill, J.S. (1874) *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific evidence.* Eighth Edition. New York: Harper. (First edition published 1843.)

Novick, L.R., & Cheng, P.W. (2004). Assessing interactive causal influence. *Psychological Review*, *111*, 455-485.

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

Perales, J.C., & Shanks, D.R. (2008). Driven by power? Probe question and presentation format effects on causal judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1482–1494.

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

Russell, B.A.W. (1912-13). On the notion of cause. *Proceedings of the Aristotelian Society*, *13*, 1-26.

Salmon, W.C. (1980). Probabilistic causality. *Pacific Philosophical Quarterly*, 61, 50-74.

Salsburg, D. (2001). *The lady tasting tea: How statistics revolutionized science in the Twentieth century.* New York: W.H. Freeman.

Sanbonmatsu, D. M., Posavac, S. S., Vanous, S., & Ho, E. A. (2005). Information search in the testing of quantified hypotheses: How "all," "most," "some," "few," and "none" hypotheses are tested. *Personality and Social Psychology Bulletin*, *31* (2), 254-266.

Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child Development, 77*(2), 427-442.

Shanks, D. R. (2004). Judging covariation and causation. In Koehler, D.J., and Harvey, N. (Eds.) *Handbook of judgment and decision making*. (pp. 220-239) Oxford: Blackwell.

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

Sloman, S. A., Barbey, A. K., & Hotaling, J. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science, 33*, 21-50.

Sloman, S.A., & Lagnado, D.A. (2005) Do we 'do'? *Cognitive Science, 29*, 5-39.

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General, 126,* 323-348.

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

Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North-Holland.

Suppes, P. (1984). *Probabilistic metaphysics*. Oxford: Basil Blackwell.

Tooley, M. (1987). *Causality: a realist approach*. Oxford: Oxford University Press.

Turnbull, W., & Slugoski, B.R. (1988). Conversational and linguistic processes in causal attribution. In Hilton, D. (Ed.) *Contemporary science and natural explanation*: *Commonsense conceptions of causality*. Brighton, Sussex: Harvester Press. Pp. 66-93.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In Fishbein, M. (Ed.) *Progress in social psychology*. Hillsdale, NJ: Erlbaum. Pp. 49-72.

White, P. A. (1995). Use of prior beliefs in the assignment of causal roles: causal powers versus regularity-based accounts. *Memory and Cognition*, 23, 243–254.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General, 136*, 82-111.

Wolff, P. & Song, G. (2003). Models of causation and causal verbs. *Cognitive Psychology, 47*, 276-332.

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

Figure captions.

Figure 1. The percentages of refutations in Experiment 1 of the forms: *A* and *not-B*, *not-A* and *B*, and *A* and *B* for the three sorts of causal relations.

Figure 2. The percentages of single and multiple refutations required for causal and enabling assertions of three sorts: universal (*in every case*), generic (no quantifier), and existential (*in some cases*) in Experiment 4.

Figure 3. The percentages of *A and not-B* and *not-A and B* refutations for causal and enabling assertions of three sorts: universal (*in every case*), generic (no quantifier), and existential (*in some cases*) in Experiment 4.

Appendix

Materials used in the four experiments. The first three items in each domain were used in

all experiments as was the fourth item in the mechanical domain (item 16). The

remaining items were generated for Experiment 4.

Biology / physiology domain	Causes	Enables	Select one of the two possibilities
1	Following this diet causes a person with this sort of metabolism to lose weight.	Following this diet enables a person with this sort of metabolism to lose weight.	A person with this sort of metabolism followed this diet and yet did not lose weight. A person with this sort of metabolism did not follow this diet and yet lost weight.
2	Regular exercise of this sort causes a person to build muscle growth.	Regular exercise of this sort enables a person to build muscle.	A person engaged in regular exercise of this sort and yet did not build muscle. A person did not engage in regular exercise of this sort and yet built muscle.
3	Having this drink causes a person with this disease to feel better.	Having this drink enables a person with this disease to feel better.	A person with this disease had this drink and yet did not feel better. A person with this disease did not have this drink and yet felt better.
4	Below zero temperatures cause this mammal to go into hibernation.	Below zero temperatures enable this mammal to go into hibernation.	Temperatures were below zero and yet this mammal did not go into hibernation. Temperatures were not below zero and yet this mammal went into hibernation.
5	Adding this	Adding this	This substance was

	substance to the soil causes this plant to produce large flowers.	substance to the soil enables this plant to produce large flowers.	added to the soil and yet the plant did not produce large flowers. This substance was not added to the soil and yet the plant produced large flowers.
6	Injection of this substance into a person with this disorder causes their blood to clot.	Injection of this substance into a person with this disorder enables their blood to clot.	A person with this disorder received an injection of this substance and yet their blood did not clot. A person with this disorder did not receive an injection of this substance and yet their blood clotted.
Natural domain			
7	Pouring this liquid onto magnesium causes it to burn.	Pouring this liquid onto magnesium enables it to burn.	This liquid was poured onto magnesium and yet it did not burn. The liquid was not poured onto magnesium and yet it burnt.
8	Contact between these two sorts of substance causes an explosion to occur.	Contact between these two sorts of substance enables an explosion to occur.	These two sorts of substance came into contact and yet no explosion occurred. These two sorts of substance did not come into contact and yet an explosion occurred.
9	Bombarding this substance with alpha rays causes it to become phosphorescent.	Bombarding this substance with alpha rays enables it to become phosphorescent.	This substance was bombarded with alpha rays and it did not become phosphorescent. This substance was not bombarded with alpha rays and yet it became phosphorescent.
10	Combining these two ingredients causes the dough to expand.	Combining these two ingredients enables the dough to expand.	These two ingredients were combined and yet the dough did not expand. These two ingredients

			were not combined and yet the dough expanded.
11	Introduction of this coil causes the cylinder to become magnetised.	Introduction of this coil enables the cylinder to become magnetised.	This coil was introduced and yet the cylinder did not become magnetised. This coil was not introduced and yet the cylinder became magnetised.
12	Heating this substance above a certain temperature causes it to turn to liquid.	Heating this substance above a certain temperature enables it to turn to liquid.	The substance was heated above this certain temperature and yet it did not turn to liquid. The substance was not heated above this certain temperature and yet it turned to liquid
Mechanical domain			
13	Tuning this kind of engine in this special way causes a reduction in its fuel consumption to occur.	Tuning this kind of engine in this special way enables a reduction in its fuel consumption to occur.	An engine of this kind was tuned in this special way and there was no reduction in its fuel consumption. An engine of this kind was not tuned in this special way and yet there was a reduction in its fuel consumption.
14	Inserting graphite rods into a nuclear reactor causes its activity to slow down.	Inserting graphite rods into a nuclear reactor enables its activity to slow down.	Graphite rods were inserted into a nuclear reactor and its activity was not slowed down. Graphite rods were not inserted into a nuclear reactor and yet its activity was slowed down.
15	Pulling this lever causes this sort of machine to start.	Pulling this lever enables this sort of machine to start.	The lever was pulled and a machine of this sort did not start. The lever was not pulled and yet a machine of this sort started.
16	Pushing the lid causes this type of	Pushing the lid enables this type of	A person pushed the lid and yet a container of

	container to open.	container to open.	this type did not open. A person did not push the lid and yet a container of this type opened.
17	Turning this dial causes steam to escape.	Turning this dial enables steam to escape.	The dial was turned and yet no steam escaped. The dial was not turned and yet steam escaped.
18	Flipping this switch causes balls to be released.	Flipping this switch enables balls to be released.	The switch was flipped and yet balls were not released. The switch was not flipped and yet balls were released.
Socio-economical domain			
19	Introducing these health care reforms causes more people to seek medical attention.	Introducing these health care reforms enables more people to seek medical attention.	These health care reforms were introduced and yet more people did not seek medical attention. These health care reforms were not introduced and yet more people sought medical attention.
20	Cutting interest rates causes derivative values to rise.	Cutting interest rates enables derivative values to rise.	Interest rates were cut and derivative values did not rise. Interest rates were not cut and yet derivative values rose.
21	This kind of a sudden change in workers' conditions causes them to earn more.	This kind of a sudden change in workers' conditions enables them to earn more.	This kind of a sudden change in workers' conditions occurred and yet they did not earn more. No sudden change of this kind in workers' conditions occurred and yet they earned more.
22	Introducing these measures causes people to	Introducing these measures enables people to	These new measures were introduced and yet people did not

	demonstrate.	demonstrate.	demonstrate. These new measures were not introduced and yet people demonstrated.
23	Implementing this policy causes a person in this tax band to buy property abroad.	Implementing this policy enables a person in this tax band to buy property abroad.	This policy was implemented and yet a person in this tax band did not buy property abroad. This policy was not implemented and yet a person in this tax band bought property abroad.
24	Participation in this programme causes a person's literacy levels to rise.	Participation in this programme enables a person's literacy levels to rise.	A person participated in this programme and yet their literacy levels did not rise. A person did not participate in this programme and yet their literacy levels rose.
Psychological domain			
25	A shock to the system of this sort causes you to have a pleasant memory of unpleasant events.	A shock to the system of this sort enables you to have a pleasant memory of unpleasant events.	A person experienced this sort of shock to the system and did not have a pleasant memory of unpleasant events. A person did not experience this sort of shock to the system and yet had a pleasant memory of unpleasant events.
26	A person's positive attitude towards you causes you to like that person.	A person's positive attitude towards you enables you to like that person.	A person had a positive attitude towards you and yet you did not like that person. A person did not have a positive attitude towards you and yet you liked that person.
27	Using this technique causes a person to experience less pain	Using this technique enables a person to experience less pain	A person used this technique and yet did not experience less pain during this type of task.

	during this type of task.	during this type of task.	A person did not use this technique and yet experienced less pain during this type of task.
28	A diagrammatic presentation of the problem causes improved performance.	A diagrammatic presentation of the problem enables improved performance.	The problem was presented diagrammatically and yet performance did not improve. The problem was not presented diagrammatically and yet performance improved.
29	Increasing group cohesiveness causes a group to perform more efficiently.	Increasing group cohesiveness enables a group to perform more efficiently.	Group cohesiveness was increased and yet a group did not perform more efficiently. Group cohesiveness was not increased and yet a group performed more efficiently.
30	Consumption of flavonoid-rich chocolate causes a person to perform better on this task.	Consumption of flavonoid-rich chocolate enables a person to perform better on this task.	A person consumed flavonoid-rich chocolate and yet did not perform better on this task. A person did not consume flavonoid-rich chocolate and yet performed better on this task.

Additional items in Experiment 3

Strong cause:

Drinking a lot of alcohol causes a person to get drunk, i.e. brings about drunkenness.

Weak cause:

Drinking a tumbler of water causes a person's thirst to be quenched, i.e. brings about the

quenching of a person's thirst.

Strong enabler:

Low atmospheric pressure enables a fog to occur, i.e. makes it possible for fog to occur.

Weak enabler:

A magnifying glass enables a person to read small text, i.e. makes it possible for a person

to read small text.