# COGNITIVE SCIENCE
## A Multidisciplinary Journal

# The Causal Structure of Utility Conditionals

## Jean-François Bonnefon,[a] Steven A. Sloman[b]

[a]*Centre National de la Recherche Scientifique and Université de Toulouse*
[b]*Department of Cognitive, Linguistic, & Psychological Sciences, Brown University*

## Abstract

The psychology of reasoning is increasingly considering agents' values and preferences, achieving greater integration with judgment and decision making, social cognition, and moral reasoning. Some of this research investigates utility conditionals, ''if $p$ then $q$'' statements where the realization of $p$ or $q$ or both is valued by some agents. Various approaches to utility conditionals share the assumption that reasoners make inferences from utility conditionals based on the comparison between the utility of $p$ and the expected utility of $q$. This article introduces a new parameter in this analysis, the underlying causal structure of the conditional. Four experiments showed that causal structure moderated utility-informed conditional reasoning. These inferences were strongly invited when the underlying structure of the conditional was causal, and significantly less so when the underlying structure of the conditional was diagnostic. This asymmetry was only observed for conditionals in which the utility of $q$ was clear, and disappeared when the utility of $q$ was unclear. Thus, an adequate account of utility-informed inferences conditional reasoning requires three components: utility, probability, and causal structure.

*Keywords:* Reasoning; Utility; Cause; Diagnostic; Conditional

## 1. The causal structure of utility conditionals

The psychology of reasoning is entering a new paradigm (Over, 2009). One pillar of this new paradigm is an affirmation that inferences depend not only on logic but also on value. Goals, valence, preferences, and utility cannot be ignored in the psychology of reasoning (Oaksford et al., 1999). A large amount of empirical work is now devoted to the inferences that people draw from premises with valued contents. A large part of this work has been conducted with conditionals, for they naturally describe actions and outcomes:

Correspondence should be sent to Jean-François Bonnefon, CLLE Maison de la recherche, 5 allées A. Machado, 31058 Toulouse Cedex 9, France. E-mail: bonnefon@univ-tlse2.fr

(1)    a.    If the CEO admits to the fraud, he will be sent to jail.
       b.    If you connect to that site, my computer will be infected by a virus.

These two examples are *utility conditionals,* ''if *p*, then *q*'' statements where *p* or *q* has some value for some agents (Bonnefon, 2009). They invite inferences that a purely epistemic approach cannot predict. For example, the CEO will not admit to the fraud, and the listener is not going to connect to the site.

Such conditionals are a hot topic in reasoning research, for they hold a promise to align the interests of the field to that of judgment and decision making, argumentation research, folk understanding of mental states (Malle, 2004), and valence effects in moral reasoning (Greene, 2007). Accordingly, a growing body of data has accumulated that addresses the way people process utility conditionals (e.g., Bonnefon, 2012; Bonnefon & Hilton, 2004; Bonnefon et al., 2012; Corner et al., 2011; Evans et al., 2008; Haigh et al., 2011; Ohm & Thompson, 2004).

Although this research has generated a broad range of new empirical findings, it has not fully connected yet to prior research on epistemic reasoning. A case in point is the problem of how causal structure interacts with preferences to generate utility-informed inferences from conditionals. Research on utility conditionals predicts how preferences generate inferences, but it is silent so far about the way the causal structure of utility conditionals can moderate these inferences. Conversely, research on epistemic reasoning has paid close attention to the way causal structure affects conditional reasoning, but it is silent (by definition and in practice) about inferences that derive from the preferences various agents may have about the propositions featured in the conditionals. Our goal in this article is to better understand how preferences and causal structure interact to generate utility-informed conditional reasoning.

## 2. Utility conditionals

Utility conditionals are ''if *p* then *q*'' conditionals such that *p* or *q* or both have some utility for some agents, for example:

(2)    a.    If I eat oysters, I become very sick.
       b.    If she makes that deal, I am ruined.
       c.    If you vote for him, he will reward you.

Such conditionals trigger inferences about what the agents are likely to do or feel, for example:

(3)    a.    The speaker will not eat oysters.
       b.    The speaker thinks she should not make that deal.
       c.    The listener is going to vote for him.

Theories of utility conditionals predict these inferences based on assumptions about the way reasoners believe others make their decisions (Bonnefon, 2011). These theories typically present three features: (a) they include a representation of utility, either in terms of goals, preferences, or costs and benefits; (b) they endow reasoners with general

Theory-of-Mind beliefs about how other people make their decisions; and (c) they predict what reasoners infer, on these bases, about the behavior and attitudes of others.

The closest precursors to theories of utility conditionals are content-specific approaches to reasoning, such as pragmatic reasoning schemas (Cheng & Holyoak, 1985) or Darwinian algorithms (Cosmides et al., 2010). These approaches introduced considerations of costs, benefits, and decision rules within the psychology of reasoning, but they differ from theories of utility conditionals in two main respects. First, they were designed to be domain-specific, only addressing subclasses of utility conditionals (e.g., permissions, hazard management, sharing rules). Second, and critically, the goals of these approaches were to explain how reasoners would detect norm violations, not how they would predict behavior. That is, they aimed at explaining how reasoners detected someone who had permission or an obligation, or someone who cheated. In contrast, theories of utility conditionals aim at explaining how reasoners predict that someone will cheat, or help, or make a sacrifice, or make any other decision that involves valued consequences.

Content-specific approaches to reasoning are more closely related to theories of utility conditionals than domain-general takes on content effects, such as the pragmatic modulation principle of Mental Model theory (Johnson-Laird & Byrne, 2002). Mental Model theory has never been satisfactorily applied to utility-informed inferences. Because it does not have a way to represent utility, nor assumptions about how agents make their decisions, Mental Model Theory does not seem to have the capacity to provide an a priori account for utility-informed inferences. Additional assumptions, though, might allow for a post hoc account. In the theory of Bonnefon (2009), assumptions about how agents make decisions take the form of *folk axioms of decision,* for example:

**Folk Axiom 1** (Self-Interested Behavior). Agents take actions that increase their own personal utility (and they do not take actions that decrease their own personal utility).

**Folk Axiom 2** (Self-Interested Attitude). Agents think that actions increasing their own personal utility should be taken by others, when others can take these actions (and they think that actions decreasing their own personal utility should not be taken by others, when others can take these actions).

These folk axioms are not meant to be descriptive of what agents do or think in real life. They aim to capture default assumptions that people make about the actions and feelings of other agents (Malle, 2004; Miller, 1999; Smedslund, 1997). Folk Axiom 1 applied to (2-a) predicts the inference ''The speaker will not eat oysters'' and Folk Axiom 2 applied to (2-b) predicts the inference ''The speaker thinks she should not make that deal.'' The broad principles captured by the folk axioms also support other approaches to utility conditionals. For example, Evans et al. (2008) and Corner et al. (2011) would explain the inference ''The speaker is not going to eat oysters'' as resulting from (a) an unfavorable comparison between the benefits of $p$ (eating oysters) and the costs of $q$ (being very sick), and (b) a sufficiently high conditional probability of $q$ given $p$.

Current approaches to utility conditionals take reasoning research beyond the realm of purely epistemic reasoning, but they have not connected yet with the research on causal structure that has had a profound impact on the psychology of reasoning. In the next section, we consider how causal structure can interact with utility in a way that is not predicted yet by theories of utility conditionals.

## 3. Causal structure

Long since the seminal paper of Cummins et al. (1991), it has been robustly established that conditional sentences of the form ''if *p* then *q*'' can be categorized as a function of the underlying causal structure that they express, and that conditional inferences track this underlying causal structure (Ali et al., 2011; Evans et al., 2007; Politzer & Bonnefon, 2006; Sloman, 2005; Sloman & Lagnado, 2005; Verbrugge et al., 2007; Weidenfeld et al., 2005).

Linguists have offered various such typologies of conditional sentences (Dancygier, 1998; Declerck & Reed, 2001; Sweetser, 1990), all of which include some form of contrast between causal and diagnostic conditionals.[1] In causal conditionals such as (4-a), *p* is a cause of *q* (in the sense that an intervention on *p* would impact on *q*), whereas in diagnostic conditionals such as (4-b), *p* is an indication that *q* is true:

(4)    a.    If the patient is infected, then she has a fever.
       b.    If the patient has a fever, then she is infected.

Epistemic inferences from conditionals such as (4-ab) track their underlying causal structure, and we predict that this underlying causal structure will also moderate inferences triggered by the utility of *p* and *q*. Consider, for example:

(5)    If he accepts this deal, then he is ruined.

All current approaches to utility-informed inferences would consider that (a) the negative utility of being ruined looms larger than the unknown utility of accepting the deal, (b) there is a significant probability of being ruined given that the deal is accepted, therefore (c) the conditional triggers the inference that the male agent will not accept the deal.

We propose that one missing element in this analysis is the causal structure of Conditional (5). Specifically, we predict that the inference is only triggered when the underlying structure of (5) is causal (e.g., accepting the deal would cause him to be ruined, because it is a very bad deal), but not when the underlying structure of (5) is diagnostic (e.g., accepting the deal would indicate that he is ruined, because the deal only appeals to desperate persons). However unfavorable the comparison between the utility of *q* and that of *p*, and however high the probability of *q* given *p*, no conclusion should be reached about whether a rational agent would do *p* when *p* is merely a symptom of *q*. Only when the realization of *p* is causally instrumental to the realization of *q*, should inferences be derived from the utilities attached to *p* and *q* and the probability of *q* given *p*.

More generally, we predict that the utility-informed inferences which were observed by prior research will be supported by utility conditionals whose underlying structure is causal, but not comparably supported by utility conditionals whose underlying structure is diagnostic.[2] Our general claim is that analyses of utility conditionals that focus on utility and conditional probability must be complemented with considerations of causal structure to make correct predictions. In this article, we offer four experiments testing various aspects of this general claim.

The goal of our first experiment is to demonstrate that inferences predicted by the folk axiom of Self-Interested Behavior (or other similar principles) are weakened when the underlying structure of the conditional is diagnostic. To do so, we will take advantage of utility conditionals whose underlying causal structure can be manipulated while their form is kept constant. We will present participants with utility conditionals such as ''If he accepts this deal then he is ruined,'' and we will record their tendency to make inferences such as ''He will accept the deal.'' We expect that participants will make these inferences when the underlying structure of the conditional is causal, but less so when the underlying structure of the conditional is diagnostic.

## 4. Experiment 1

### 4.1. Method

Participants were 41 students at the University of Toulouse (seven men, mean age $=$ 21), who were individually approached on campus by a research assistant. They all read the following introduction to the experiment:[3]

> We are about to present you with eight situations, which are all briefly described in a couple of sentences. Each situation features a different character, denoted by ''he'' or ''she.'' For example:
> *If she eats oysters, then she is very sick. (Because she is allergic to oysters.)*
> We ask you, for each situation, whether you expect the character to do something or not. For example:
> *If she going to eat oysters?*
> Of course, there is no correct or incorrect answer. We wish to know what spontaneously comes to your mind. To answer, you must check one box on a scale from −5 to +5. Zero means that you have no idea at all in one direction or the other.

This text was followed by an example of the 11-point response scale, anchored at *certainly not* and *certainly*, whose points were labeled from −5 to +5.

The eight conditionals (four with a positive consequent, four with a negative consequent) were the same for all participants. Four of these were glossed (in parentheses) to cue a causal interpretation and the other four were glossed to cue a diagnostic interpretation. Two versions of the questionnaire were constructed so that the conditionals which received a causal gloss in one version received a diagnostic gloss in the other. In addition, the order in which the conditionals appeared was counterbalanced across questionnaires.

To illustrate (see Appendix A for a complete list), here is an example of one conditional together with its causal gloss (6-a) and together with its diagnostic gloss (6-b):

(6)    a.    If he buys this house, then he is rich. (Because he can resell it for twice the price.[4]) Is he going to buy this house?

         b.    If he buys this house, then he is rich. (Because it costs a fortune.) Is he going to buy this house?

We expect a polarization of responses in the causal condition, as compared to the diagnostic condition. That is, responses should be more positive for causal conditionals with positive consequences and more negative for causal conditionals with negative consequences.

## 4.2. Results

In order to construct an index of utility-informed reasoning, we changed the sign of the responses to the four problems featuring a negative-utility consequent. That way, higher ratings always denote higher endorsement of the utility-informed conclusion. If participants take into account the causal structure suggested by the gloss, then their index of utility-informed reasoning should be higher for causal glosses than for diagnostic glosses.[5]

As predicted, the average index of utility-informed reasoning was 2.5 (MSE = 0.3) for causal glosses, and only 0.9 (MSE = 0.2) for diagnostic glosses, $t(40) = 3.6$, $p = .001$. The 95% confidence interval for the difference between the two values of the index is 0.7–2.4, for a large effect size $d = 0.94$. The comparison across items confirmed that the index of utility-informed reasoning was greater for causal glosses than for diagnostic glosses, $t(7) = 3.8$, $p = .007$. These initial findings suggest that reasoners derived inferences from utility conditionals whose underlying structure was that $p$ caused $q$, but they were less ready to do so from utility conditionals whose underlying structure was that $p$ was diagnostic of $q$.

The aim of Experiment 2 is to consolidate this result by introducing a more controlled manipulation of causal structure, and by ruling out one alternate explanation. We found that the likelihood of $p$ was different depending on whether the utility conditional had a causal or diagnostic structure. The possibility we need to rule out is that our manipulation of structure might have cued people to imagine scenarios where $q$ itself was more or less likely, and that this effect might have propagated to the likelihood of $p$.

To illustrate, consider the following example:

(7)  a.  If she moves to Paris, then she no longer works for her company. (Because her company will disapprove.) Is she going to move to Paris?
     b.  If she moves to Paris, then she no longer works for her company. (Because it is a good place for her to look for another job.) Is she going to move to Paris?

In this example, a case can be made that (7-a) makes it sound more likely than (7-b) that she still works for the company, that is, $q$ seems to have lesser likelihood in (7-a) than in (7-b). As a consequence, $p$ might be given lesser likelihood in (7-a) than in (7-b) simply because $q$ is less probable. In other words, our manipulation might have primarily affected the likelihood of $q$ and only secondarily by propagation the likelihood of $p$.

To assess this explanation, we need reasoners to rate both the likelihood of $p$ and the likelihood of $q$. If our original explanation is correct, we should find again a strong effect on the likelihood of $p$, and no comparable effect on the likelihood of $q$.[6] If the alternative explanation is correct that the primary effect of our manipulation is on $q$, we should again find a strong effect on the likelihood of $p$, but an even stronger effect on the likelihood of $q$.

## 5. Experiment 2

In our first experiment, we used a variety of glosses to cue the causal structure of our conditional. Experiment 2 introduces a more controlled manipulation of causal structure. We also separate the causal explanation from the conditional sentence. As a result, the two causal structure conditions use identical conditional sentences.

### 5.1. Method

Participants were 38 students at the University of Toulouse (12 men, mean age = 25), who were individually approached on campus by a research assistant. The method and materials were similar to that of Experiment 1, except for two changes. The first change related to the manipulation of causal structure. In Experiment 2, the causal structure was first spelled out in a systematic way, before it was paraphrased with the conditional. For example:

(8)  a.  The fact that he buys this house would make him rich as a result. In other words: If he buys this house, then he is rich.
  b.  The fact that he buys this house would indicate that he is rich. In other words: If he buys this house, then he is rich.

In all problems, the conditionals in the causal condition were preceded by a sentence of the form ''The fact that $p$ would make $q$ true as a result,'' and the conditionals in the diagnostic condition were preceded by a sentence of the form ''The fact that $p$ would indicate that $q$ is true.''

The second change was that participants rated the likelihood of $q$ in addition to the likelihood of $p$ on a similar 11-point scale anchored at *Certainly not* and *Certainly,* whose points were labeled from −5 to +5.

### 5.2. Results

The ratings participants gave about $p$ almost exactly replicated the findings of Experiment 1: Utility-informed inferences about $p$ were strong in the causal condition ($M = 2.3$, MSE = 0.3) and much weaker in the diagnostic condition ($M = 1.0$, MSE = 0.2). The manipulation of causal structure only had a small effect on $q$, as ratings in the causal condition ($M = 1.0$, MSE = 0.2) were very similar to ratings in the diagnostic condition ($M = 0.6$, MSE = 0.2).

We conducted a repeated-measure analysis of variance, where the object of the rating ($p$ or $q$) and the condition were entered as predictors. The analysis revealed a main effect reflecting the fact that $p$ received greater ratings than $q$, $F(1,37) = 15$, $p < .001$, $\eta_p^2 = 0.29$; and a main effect of condition, $F(1,37) = 10$, $p = .003$, $\eta_p^2 = 0.22$. Critically, the analysis detected the interaction effect reflecting the fact that the experimental condition had a larger impact on $p$ than on $q$, $F(1,37) = 13$, $p < .001$, $\eta_p^2 = 0.26$. Specifically, causal structure had a strong effect on the likelihood of $p$, $t(37) = 4.1$, $p < .001$, but no comparable effect on the likelihood of $q$, $t(37) = 1.7$, $p = .09$. This

was consistent with the results of a comparison across items, which detected a significant effect of causal structure on $p$-ratings, $t(7) = 3.0$, $p = .02$, but no comparable effect on $q$-ratings, $t(7) = 2.2$, $p = .07$.

Experiment 2 replicated the main finding of Experiment 1 using a more controlled manipulation: Participants took into account the causal structure of utility conditionals when engaging in utility-informed reasoning. Specifically, they drew inferences as predicted by the folk axiom of Self-Interested Behavior, but only when the underlying structure of the conditional was that $p$ caused $q$. The aim of Experiment 3 is to generalize this finding to another inference, based on the folk axiom of Self-Interested Attitude (or other similar principles).

## 6. Experiment 3

In Experiment 3, we extend our investigations to inferences about what agents would think, based on the folk axiom of Self-Interested Attitude. We use statements such as:

(9)    If she buys a house, then he is rich.

We predict that the inference ''He thinks she should buy the house'' will be triggered when the underlying structure of (9) is causal, but not when the underlying structure of (9) is diagnostic.

### 6.1. Method

Participants were 40 students at the University of Toulouse (8 men, mean age = 23), who were individually approached on campus by a research assistant. They all read the following introduction to the experiment:

> We are about to present you with eight situations, which are all briefly described in a couple of sentences. Each situation features two characters, denoted by ''he'' and ''she.'' For example:
> *If she files for divorce, he will lose everything he has.*
>  We ask you, for each situation, whether you expect that ''he'' thinks ''she'' should or should not do something. For example:
> *Does he think she should file for divorce?*
>  Of course, there is no correct or incorrect answer. We wish to know what spontaneously comes to your mind. To answer, you must check one box on a scale from −5 to +5. Zero means that you have no idea at all, in one direction or the other.

This text was followed by an example of the 11-point response scale, anchored at *certainly not* and *certainly*, whose points were labeled from −5 to +5.

The eight conditionals were the same for all participants. Four of these were introduced as paraphrases of an underlying causal structure, and the other four were introduced as paraphrases of an underlying diagnostic structure (using the same method as in Experiment 2). Two versions of the questionnaire were constructed, so that the conditionals which paraphrased an underlying causal structure in one version paraphrased an underlying diagnostic structure in the other version. In addition, the order in which the conditionals appeared was counterbalanced across questionnaires.

To illustrate (see Appendix B for a complete list), here is an example of one conditional in the two conditions:

(10)  a.  The fact that she buys a house would make him rich as a result. In other words: If she buys a house, then he is rich. Does he think she should buy a house?
      b.  The fact that she buys a house would indicate that he is rich. In other words: If she buys a house, then he is rich. Does he think she should buy a house?

## 6.2. Results

As anticipated, the average index of utility-informed reasoning was 2.9 (MSE = 0.3) for causal glosses, and only 1.7 (MSE = 0.3) for diagnostic glosses, $t(39) = 3.5, p = .001$. The 95% confidence interval for the difference between the two values of the index is 0.5–1.9, for an effect size $d = 0.55$. The comparison across items confirmed that the index of utility-informed reasoning was greater for causal glosses than for diagnostic glosses, $t(7) = 5.7$, $p = .001$. This result further establishes that reasoners take into account the causal structure of utility conditionals when engaging in utility-informed reasoning. Once more, they derived inferences from utility conditionals whose underlying structure was that $p$ caused $q$, but they were less ready to do so from utility conditionals whose underlying structure was that $p$ was diagnostic of $q$.

## 7. Experiment 4

In this final experiment, we aim to demonstrate that the asymmetries observed in Experiments 1–3 occur only for utility conditionals, and thus cannot be explained by purely epistemic models of reasoning. To do this, we contrast conditionals in which $q$ clearly has value to conditionals in which the value of $q$ is unclear. We expect that causal structure will affect utility-informed inferences from the former sort of conditionals, but not from the latter sort. In addition, Experiment 4 extends our investigations using an expression of causal structure that includes tense, providing for more natural sentences.

### 7.1. Method

Participants were 19 men and 34 women recruited through the Mechanical Turk platform (mean age = 33, $SD$ = 11). They read an introduction to the study similar to that used in Experiment 3 and were presented with 32 scenarios presented in randomized order. The 32

scenarios consisted of four different versions of the eight scenarios used in Experiment 3, following a $2 \times 2$ within-participant design manipulating the utility of $q$ (clear vs. unclear) and the underlying causal structure of the conditional (causal vs. diagnostic). Here is an example of a single scenario in all four conditions:

(11)  a.  Her accepting the deal would cause him to be ruined. In other words, if she accepts the deal, then he will be ruined. Does he think she should accept the deal? *causal-clear utility*

b.  Her accepting the deal would indicate that he is ruined. In other words, if she accepts the deal, then he must be ruined. Does he think she should accept the deal? *diagnostic-clear utility*

c.  Her accepting the deal would cause him to have different prospects. In other words, if she accepts the deal, then he will have different prospects. Does he think she should accept the deal? *causal-unclear utility*

d.  Her accepting the deal would indicate that he has different prospects. In other words, if she accepts the deal, then he must have different prospects. Does he think she should accept the deal? *diagnostic-unclear utility*

As illustrated in (11-ad), the manipulation of causal structure involved a change from ''will'' (causal) to ''must'' (diagnostic), and unclear utility was obtained by mentioning that the situation of the agent would change, without specifying the direction of that change. As in previous experiments, participants used an 11-point response scale, anchored at *certainly not* and *certainly*, whose points were labeled from −5 to +5.

*7.2. Results*

In line with the results of Experiments 1–3, causal structure impacted the utility-informed reasoning index for scenarios where utility was clear. The average index was 3.7 (MSE = 0.2) for causal statements, and only 2.8 (MSE = 0.2) for diagnostic statements, $t(52) = 5.1$, $p < .001$. The 95% confidence interval for the difference between the two values of the index was 0.5–1.2, for an effect size $d = 0.53$.[7]

Causal structure did not have any detectable impact, however, for scenarios where utility was unclear. The average index was −0.2 (MSE = 0.1) for causal statements, and −0.1 (MSE = 0.1) for diagnostic statements, $t(52) = -0.8$, $p = .42$. The 95% confidence interval for the difference between the two values of the index was −0.3−+0.1, for an effect size $d = 0.12$. A repeated-measure analysis of variance confirmed a statistically significant interaction between our two variables, $F(1,52) = 26$, $p < .001$.

The effect of causal structure on inferences from utility conditionals was thus confirmed again, using yet another manipulation of causal structure. This effect was shown to be specific to utility conditionals, and not observed for conditionals in which the utility of $q$ was unclear.

# 8. General discussion

The psychology of reasoning is devoting increasing attention to inferences stemming from the preferences that reasoners attribute to various agents, and it is thus achieving

greater integration with other subfields such as judgment and decision making, social cognition, and moral reasoning.

A large share of this new research investigated utility conditionals, ''if *p* then *q*'' statements where the realization of *p* or *q* or both is valued by some agents. Various approaches coexist to predict what reasoners infer about the actions and feelings of the interested agents from these conditionals for example, (e.g., Bonnefon, 2009; Corner et al., 2011; Evans et al., 2008). At the core of these approaches is a common assumption that reasoners base their inferences on a comparison between the utility of *p* and the expected utility of *q*. Thus, all these approaches would make the following prediction: When an action *p* in ''if *p* then *q*'' has a low positive utility for some agent *x*, *q* has a high cost for this same agent, and the probability of *q* given *p* is high, then the conditional triggers an inference that agent *x* is not going to take action *p*. In this article, we have argued for the introduction of a new parameter in the analysis of inferences from utility conditionals, that of the underlying causal structure of the conditional.

In four experiments, we showed that causal structure moderates inferences from utility conditionals. These experiments investigated inferences about actions (Experiments 1 and 2) and thoughts (Experiments 3 and 4) and used three different manipulations of causal structure. In all cases, utility-informed inferences were strongly invited when the underlying structure of the conditional was causal and significantly less so when the underlying structure of the conditional was diagnostic. This asymmetry was only observed for conditionals in which the utility of *q* was clear, and disappeared for conditionals in which the utility of *q* was unclear. The lesson from these four experiments is that an adequate account of inferences from utility conditionals requires three components: utility, probability, and causal structure. Without considerations of utility and probability, we cannot account for the utility-informed inferences observed in prior research, and robustly replicated in the causal condition of the current experiments. Without considerations of causal structure, we cannot explain why utility-informed inferences disappear in the diagnostic condition of Experiments 1–3.

One argument, though, might be considered against the conclusion that causal structure is required in models of utility-informed inferences: Mere temporal order could explain the findings; therefore, considerations of causal structure are not required. Our causal and diagnostic conditionals differed on one dimension besides causal structure that of temporality. In causal conditionals, *p* would occur before *q*, whereas in diagnostic conditionals *p* would occur after *q*. Thus, one might wonder whether temporal order alone might be responsible for the disappearance of utility-informed inferences without the need to consider causal structure. It is of course impossible to conceive realistic scenarios where causation would go backward in time, but the next best thing is to show that utility-informed inferences require a causal link between *p* and *q* in addition to mere temporal order. Consider the following example. The first author of this article was born in early November and enjoys celebrating his birthday. He might say that ''If children come trick-or-treating, then my birthday is next week,'' where there is a temporal, non-causal relation between the action *p* and the desirable state *q*. We do not infer in that case that he wants children to come trick-or-treating. Temporal order without a causal link does not support utility-informed inferences. Mere temporal order alone is thus unlikely to account for our findings.

It thus seems necessary to explicitly incorporate considerations of causal structure in current models of utility conditionals. Arguably, considerations of causal structure are already informally or implicitly present in these approaches. For example, Bonnefon (2009, p. 891) at some point refers to the utility of $q$ as ''the long-term consequences of action $p$ brought about by virtue of producing $q$;'' Corner et al. (2011, p. 136) once refer to the utility of $q$ as the ''disutility of the action to which [$p$] could lead;'' and Evans et al. (2008, p. 112) suggest that the epistemic mental model that reasoners construct of the link between $p$ and $q$ could represent causal relations, in addition to the representation of probability on which the paper focused. These considerations of causal structure are typically mentioned in passing, though, and not formally implemented in the models offered by all these authors. As a conclusion to this article, we suggest how considerations of causal structure could be formally implemented in one of these models, that of Bonnefon (2009).

In its current version, the theory predicts inferences by (a) turning utility conditionals into a grid-like representation, and (b) detecting within this grid configurations that trigger folk axiom of decisions. In the first stage, a conditional ''if $p$, then $q$'' is transformed into a utility grid of the form:

$$\left\{ \begin{array}{llll} \text{If} & x & u & y \\ \text{Then} & x' & u' & y' \end{array} \right\}.$$

The first row of the grid contains the information related to the *if*-clause of the conditional. That is, it displays the agent $x$ (left column) who can potentially take action $p$, and the utility $u$ (central column) that this action would have for a given agent $y$ (right column). The second row of the grid contains the corresponding information with respect to the *then*-clause of the conditional. In a second stage, the model looks for configurations that match situations described by the folk axioms of decisions. For example, one configuration that triggers the folk axiom of Self-Interested Behavior is

$$\left\{ \begin{array}{llll} \text{If} & x & 0 & \bullet \\ \text{Then} & \bullet & - & x \end{array} \right\},$$

where the black dot stands for any legitimate value of the parameter. This configuration triggers the inference that agent $x$ is not going to take action $p$, as its marginal utility would not compensate for the negative utility of $q$. This two-stage model aptly captures previous findings, but it fails to account for the effect of causal structure we have explored in the current article, because causal structure is never formally considered in the model. The most natural way to formally include causal structure in the model is to label the rows of the utility grid by causal role, rather than by syntactic clause, with the constraint that the first row in the grid must correspond to that of $p$ or $q$ that causes the other. When $p$ is the cause of $q$, the utility grid stays the same; the ''if'' and ''then'' labels simply become ''cause'' and ''effect'' labels:

Table 1
Confidence (mean and SE) in the conclusion $p$ as a function of the valence of $q$ and the causal structure of ''if $p$, then $q$,'' in the four experiments. Negative values indicate confidence that $q$ is false

| | Causal | | Diagnostic | |
|---|---|---|---|---|
| | Positive | Negative | Positive | Negative |
| Exp. 1 | + 3.3 (0.3) | −1.7 (0.3) | + 2.1 (0.3) | + 0.2 (0.3) |
| Exp. 2 | + 3.0 (0.3) | −1.8 (0.2) | + 1.4 (0.4) | −0.2 (0.4) |
| Exp. 3 | + 1.9 (0.4) | −3.8 (0.2) | + 1.9 (0.4) | −1.8 (0.5) |
| Exp. 4 | + 3.5 (0.2) | −3.8 (0.2) | + 3.0 (0.2) | −2.6 (0.3) |

$$\left\{ \begin{array}{cccc} \text{Cause} & x & u & y \\ \text{Effect} & x' & u' & y' \end{array} \right\}.$$

However, when $p$ is *diagnostic* of $q$, the two rows in the grid are switched, for the row that corresponds to $q$ (the cause) now moves up to the top:

$$\left\{ \begin{array}{cccc} \text{Cause} & x' & u' & y' \\ \text{Effect} & x & u & y \end{array} \right\}.$$

It can be shown that this minimal change to the theory preserves its prior explanatory power (it still make all the correct predictions it made about causal utility conditionals), while allowing it to account for our new findings (it no longer makes incorrect predictions about diagnostic utility conditionals). It is likely that other approaches can be fixed by similar changes that would, for example, move their focus from pure probability calculus to causal calculus. Such changes are clearly necessary in light of our findings: Uncertainty, utility, and causal structure are all required in a theory of utility-informed conditional reasoning.

## Notes

1. A third common category is that of *biscuit* or *relevance* conditionals, such as ''If you are thirsty, there is beer in the fridge,'' or ''If you like rugby, there is a game on TV now.'' Interestingly, neither $p$ nor $q$ in these conditionals can be said to cause one another. Their underlying causal structure has been argued to be that $p$ causes the relevance of asserting that $q$ (Bonnefon & Politzer, 2011; DeRose & Grandy, 1999; Siegel, 2006).
2. Because we take causal relations as primitives, the current article will not delve into more elaborate analyses of the meaning of causal verbs and the distinctions between causal, enabling, and prevention relations. A fuller analysis of them appears in, for example, Goldvarg and Johnson-Laird (2001) or the response to it by Sloman et al. (2009).

3. Experiments 1–3 were conducted in French. The original French version of the materials can be obtained from the corresponding author.

4. Note that the causal gloss is a separate sentence, which is not perfect French (or English) grammar. This was done to keep the conditional sentences identical between structure conditions. Experiments 2 and 3 will introduce another manipulation of causal structure that will both satisfy grammar and keep the conditional sentences identical.

5. For the sake of completeness, Table 1 displays the ratings broken down by the utility of $q$, in the four experiments. The utility-informed reasoning index is obtained by averaging the confidence in $p$ from positive conditionals, and the opposite of the confidence in $p$ from negative conditionals.

6. We might also expect a weaker effect on the likelihood of $q$, if the likelihood of $p$ propagates to that of $q$ by basic conditional reasoning.

7. Note that for the first time in our series of experiments, we observe some clear signs of utility-informed reasoning with diagnostic conditionals. We suspect that this might be due to the ambiguity of the modal ''must'' that is used to indicate diagnostic structure in Experiment 4. For example, ''if she quits her job, then he must be promoted'' might suggest something else than a diagnostic structure: that there is an obligation to promote him if she quits her job; or that promoting him is a necessary condition for her to quit her job. In both interpretations, the male agent would benefit from her quitting her job, which would trigger utility-informed reasoning predicted by the folk axiom of self-interested attitude. However, that the difference between the causal and diagnostic condition remains highly significant.

## Acknowledgments

## References

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

Bonnefon, J. F. (2009). A theory of utility conditionals: Paralogical reasoning from decision-theoretic leakage. *Psychological Review*, *116*, 888–907.

Bonnefon, J. F. (2011). Norms for reasoning about decisions. *Behavioral and Brain Sciences*, *34*, 249–250.

Bonnefon, J. F. (2012). Utility conditionals as consequential arguments: A random sampling experiment. *Thinking and Reasoning*, *18*, 379–393.

Bonnefon, J. F., Girotto, V., & Legrenzi, P. (2012). The psychology of reasoning about preferences and unconsequential decisions. *Synthese*, *187*, 27–41.

Bonnefon, J. F., & Hilton, D. J. (2004). Consequential conditionals: Invited and suppressed inferences from valued outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 28–37.

Bonnefon, J. F., & Politzer, G. (2011). Pragmatics, mental models, and one paradox of the material conditional. *Mind and Language*, *26*, 141–155.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Corner, A., Hahn, U., & Oaksford, M. (2011). The psychological mechanism of the slippery slope argument. *Journal of Memory and Language*, *64*, 153–170.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Science of the USA*, *107*, 9007–9014.

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

Dancygier, B. (1998). *Conditionals and prediction. Time, knowledge, and causation in conditional constructions*. Cambridge, England: Cambridge University Press.

Declerck, R., & Reed, S. (2001). *Conditionals. A comprehensive empirical analysis*. Berlin: Mouton de Gruyter.

DeRose, K., & Grandy, R. E. (1999). Conditional assertions and ''biscuit'' conditionals. *Noûs*, *33*, 405–420.

Evans, J. S. B. T., Handley, S. J., Hadjichristidis, C., Thompson, V., Over, D. E., & Bennett, S. (2007). On the basis of belief in causal and diagnostic conditionals. *Quarterly Journal of Experimental Psychology*, *60*, 635–643.

Evans, J. S. B. T., Neilens, H., Handley, S. J., & Over, D. E. (2008). When can we say ''if''? *Cognition*, *108*, 100–116.

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

Greene, J. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences*, *17*, 322–323.

Haigh, M., Stewart, A. J., Wood, J. S., & Connell, L. (2011). Conditional advice and inducements: Are readers sensitive to implicit speech acts during comprehension? *Acta Psychologica*, *136*, 419–424.

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

Malle, B. F. (Ed.). (2004). *How the mind explains behavior*. Cambridge, MA: MIT Press.

Miller, D. T. (1999). The norm of self-interest. *American Psychologist*, *54*, 1053–1060.

Oaksford, M., Chater, N., & Grainger, B. (1999). Probabilistic effects in data selection. *Thinking and Reasoning*, *5*, 193–243.

Ohm, E., & Thompson, V. (2004). Everyday reasoning with inducements and advice. *Thinking and Reasoning*, *10*, 241–272.

Over, D. E. (2009). New paradigm psychology of reasoning. *Thinking and Reasoning*, *15*, 431–438.

Politzer, G., & Bonnefon, J. F. (2006). Two varieties of conditionals and two kinds of defeaters help reveal two fundamental types of reasoning. *Mind and Language*, *21*, 484–503.

Siegel, M. E. A. (2006). Biscuit conditionals: Quantification over potential literal acts. *Linguistics and Philosophy*, *29*, 167–203.

Sloman, S. A. (2005). *Causal models: how people think about the world and its alternatives*. Oxford, England: Oxford University Press.

Sloman, S. A., Barbey, A. K., & Hotaling, J. M. (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.

Smedslund, J. (1997). *The structure of psychological common sense*. Mahwah NJ: Erlbaum.

Sweetser, E. E. (1990). *From etymology to pragmatics*. Cambridge: Cambridge University Press.

Verbrugge, S., Dieussaert, K., Schaeken, W., Smessaert, H., & Van Belle, W. (2007). Pronounced inferences: A study on inferential conditionals. *Thinking and Reasoning*, *13*, 105–133.

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

### Appendix A: Materials, Experiment 1

The two glosses are given between parentheses, first the causal gloss, then the diagnostic gloss. Participants saw only one of these.

1. If he quarrels with that person, then he spends time at the hospital. (Because that person is very dangerous / Because that person works at the hospital.) Is he going to quarrel with that person?
2. If she wears her dark suit, then she feels bad. (Because it does not fit / Because she is going to a funeral.) Is she going to wear her dark suit?
3. If she makes that presentation, then she is on the fast track to success. (Because it is a career booster / Because it is a meeting for senior executives.) Is she going to make that presentation?
4. If she moves to Paris, then she no longer works for her company. (Because her company will disapprove / Because it is a good place for her to look for another job.) Is she going to move to Paris?
5. If he buys this house, then he is rich. (Because he can resell it for twice the price / Because it costs a fortune.) Is he going to buy this house?
6. If he takes that drug, then he certainly recovers. (Because it is a very effective cure / Because it is given in the last stage of recovery.) Is he going to take that drug?
7. If he accepts this deal, then he is ruined. (Because it is a very bad deal / Because the deal only appeals to desperate persons.) Is he going to accept this deal?
8. If she goes to the opening, then she is famous. (Because it will give her a lot of publicity / Because everyone there is famous.) Is she going to go to the opening?

### Appendix B: Materials, Experiment 3

1. The fact that she buys a house [would make him rich as a result / would indicate that he is rich]. In other words: If she buys a house, then he is rich. Does he think she should buy a house?
2. The fact that she starts a business [would make him unemployed as a result / would indicate that he is unemployed]. In other words: If she starts a business, then he is unemployed. Does he think she should start a business?
3. The fact that she quits her job [would get him promoted as a result / would indicate that he is promoted]. In other words: If she quits her job, then he is promoted. Does he think she should quit her job?
4. The fact that she comes to the house [would make him depressed as a result / would indicate that he is depressed]. In other words: If she comes to the house, then he is depressed. Does he think she should come to the house?

5. The fact that she meets journalists [would make him famous as a result / would indicate that he is famous]. In other words: If she meets journalists, then he is famous. Does he think she should meet journalists?

6. The fact that she accepts the deal [would ruin him as a result / would indicate that he is ruined]. In other words: If she accepts the deal, then he is ruined. Does he think she should accept the deal?

7. The fact that she throws a party [would make him popular as a result / would indicate that he is popular]. In other words: If she throws a party, then he is popular. Does he think she should throw a party?

8. The fact that she feeds him grapes [would make him sick as a result / would indicate that he is sick]. In other words: If she feeds him grapes, then he is sick. Does he think she should feed him grapes?