Conditionals conflict with their denials

Sangeet Khemlani[1], Isabel Orenes[2], Sergio Moreno-Ríos[3], and Orlando Espino[4]
sangeet.khemlani@nrl.navy.mil, iorenes@psi.uned.es, semoreno@ugr.es, oespinom@ull.edu.es
[1] Navy Center for Applied Research in Artificial Intelligence, U.S. Naval Research Laboratory, USA
[2] Department of Basic Psychology, National Distance Education University, Spain
[3] Department of Developmental and Educational Psychology, University of Granada, Spain
[4] Department of Cognitive, Social, and Organisational Psychology, University of La Laguna, Spain

Abstract

What does it mean to deny the conditional statement, *if you steal an apple, you go to jail*? One theory argues that, because conditionals are probabilistic, their denials are too. And so the conditional probability, P(*not-jail* | *steal-apple*), best describes a conditional denial. Another theory argues that conditional denials concern possibilities, i.e., they activate imagined situations in which you cheat on your taxes but don't go to jail. The two accounts make diverging predictions: only the latter predicts that people should assess conditionals and their denials as mutually inconsistent. Two experiments corroborate the possibility-based account: the studies show that both in implicit and explicit evaluations of consistency, conditional denials conflict with the conditionals they deny.

Keywords: conditionals; negation; denial; mental models; probability logic

Introduction

Stealing food is wrong, and yet it rarely results in jailtime. Hence, it is reasonable to deny the conditional, *if you steal an apple, you go to jail.* What do such denials mean? One answer comes from classical systems of logic which treat conditional assertions as *material implications*. Material implications stipulate the logical properties of conditional claims of the form *if A then B;* they are true in every situation other than when the *if*-clause is true and the *then*-clause is false (Jeffrey, 1981, Ch. 4): on this account, the statement *if you steal an apple, you go to jail* is false only when you steal an apple and don't go to jail – in all other cases, the conditional is true. Some theorists accept that natural language conditionals are akin to material implications (e.g., Grice, 1989). But contemporary cognitive scientists question this equivalence because it permits implausible, "paradoxical" inferences (see, e.g., Cooper, 1968; Gazdar, 1979 for logical analyses of such inferences and Orenes & Johnson-Laird, 2012 for empirical studies on them). For instance, it treats the following inference as valid:

B.
Therefore if A then B.

That is because there is no situation in which *B* is true but *if A then B* false. The only situation that makes the conditional false is when *A* is true and *B* is false. The counterintuitive nature of the inference is transparent in this example:

Arla stole an apple.
Therefore, if Arla was amused, then she stole an apple.

People know that amusement and apple theft have little to do with one another, so the latter conclusion is implausible despite the fact that material implication guarantees its validity. An analogous criticism concerns the denials of conditionals. Consider this inference:

It's not the case that if you stole an apple, then you went to jail.
Therefore, you stole an apple.

The conclusion defies common sense: denials of *if-then* statements should not imply the truth of the *if*-clauses. Yet material implications render the above inference valid, because the only situation that makes denials of a conditional true is when *A* is true and *B* is false. So, material conditionals treat as valid any argument of the following form:

It's not the case that if A then B.
Therefore, A.

These counterintuitive implications led cognitive scientists to the widespread rejection of the material conditional (e.g., Byrne & Johnson-Laird, 2020; Evans, Handley, & Over, 2003; Kleiter, Fugard, & Pfeifer, 2018): despite their necessity in certain formal systems of symbolic logic, material conditionals cannot underlie natural language conditionals. Yet they disagree on how people represent conditional denials. Debates surrounding them have not reached any consensus (for a review, see Nickerson, 2015), but they have led to two prominent theoretical accounts: one states that conditionals describe probabilities (Baratgin et al., 2015; Douven et al., 2018; Evans & Over, 2004; Evans et al., 2003; Kleiter et al., 2018) and the other argues that they concern what's possible (Espino, Santamaría, & Byrne, 2009; Johnson-Laird & Byrne, 2002; Goodwin & Johnson-Laird, 2018; Quelhas, Rasga, & Johnson-Laird, 2018).

This paper begins by summarizing the central claims of these two cognitive theories of conditionals. It examines what they imply about how reasoners should deny conditional statements, and it outlines the predictions that tease apart the probability- and possibility-based accounts. It describes two experimental tests of the prediction, both of which support the possibility-based theory. The paper concludes by describing the mental processes implicated in the denials of conditionals.

Probabilities and possibilities

Probabilistic accounts of reasoning posit that humans reason, not in terms of certainty, but in degrees of belief (Oaksford & Chater, 2009). These degrees are best described by appealing to the norms set by the probability calculus (Evans & Over, 2013). One prominent probabilistic account of conditional reasoning is the suppositional theory (Evans & Over, 2004; Evans et al., 2003), which argues that people interpret *if A then B* using the "Ramsey test" (Ramsey, 1929/1990): they add *A* to their set of beliefs, adjust those beliefs to accommodate *A,* then assess the probability of *B*. The result is formally equivalent to treating *if A then B* as a conditional probability, i.e., *if A then B* = P(*B* | *A*), which describes the degree of belief in *B* after assuming that *A* is true. Studies show that under the right circumstances, participants who interpret *if you steal an apple, you go to jail* assign probabilities to various conjunctive combinations of the *if*- and *then*-clauses in a manner that coheres with the equation (Evans, Handley, & Over, 2003; Oberauer & Wilhelm, 2003; López-Astorga, Ragni, & Johnson-Laird, 2021). And the equation in turn serves as a lynchpin for many probabilistic accounts of conditional reasoning (for a review and debate, see Johnson-Laird, Khemlani, & Goodwin, 2015; Baratgin et al., 2015).

Probability theorists likewise concur on how people should interpret the negation of a conditional: reasoners should deny conditionals by assuming *A* is true and assessing the probability of *not-B*, i.e.: P(*not-B* | *A*) (see Handley, Evans, & Thompson, 2006, p. 559). For instance, Handley et al. note that "people understand the negation of 'if p then q' as equivalent to the conditional 'if p then not q'" (p. 569), and Pfeifer and Tulkki (2017, p. 128) make the same argument, but extend the hypothesis further to anticipate that people may interpret *if A then B* using wide-scope negation, e.g., *not*(*if A then B*). Such theories are unique in that they yield a clear and explicit prediction: a conditional and its negation can both be true – they can both have non-zero probabilities – at the same time. More formally, whenever P(*not-B* | *A*) < 1, then P(*B* | *A*) > 0. The result is a consequence of probability theory and the following identity that relates the two conditional probabilities:

P(*not-B* | *A*) = 1 - P(*B* | *A*)

Since the two probabilities must sum to 1, the only time P(*not-B* | *A*) = 0 is when P(*B* | *A*) is certain. To illustrate this situation, consider these two statements:

It's 50% certain that Arla stole an apple.	P(*A*) = 0.5
It's 50% certain Arla did not steal an apple.	P(*not-A*) = 0.5

The two statements aren't true at the same time; they're not false either; and it's not the case that one is true and the other is false. The appropriate way to conceive of them is that they are both *consistent* with one another. Probabilistic theories argue that everyday reasoning with conditionals is uncertain by default. They argue that the events underlying those possibilities (i.e., the event in which *Arla stole an apple* and the event in which *Arla did not steal an apple*) can be modeled as probabilities: P(*Arla-stole*) and P(*Arla-didn't-steal*): the two events are inconsistent with one another because no situation can exist in which both are true; but their probabilities are consistent because both probabilities can be non-zero at the same time. Hence, probabilistic theories make the following prediction: *People should judge that conditionals can have non-zero probability when their negations have non-zero probability, and vice versa*.

An alternative account posits that conditionals concern possibilities, not probabilities (Byrne & Espino, 2021; Johnson-Laird & Byrne, 2002; Khemlani, Byrne, & Johnson-Laird, 2018). So do their negations (Espino & Byrne, 2012; Khemlani, Orenes, & Johnson-Laird, 2012, 2014). The theory is akin to many philosophical and semantic treatments of conditionals that concern "possible worlds", but philosophical treatments often neglect to place constraints on the size of such worlds. As Partee (1979) notes, size matters: possibilities need to fit inside reasoners' minds. Hence, these theorists argue that people reason based on *mental* possibilities – mental models. We therefore refer to their account as the "model" theory (inspired by, but distinct from, the model theory in symbolic logic).

By default, people interpret *if you steal an apple then you go to jail* as referring to the possibility exemplified in this diagram:

```
steal-apple    jail
```

The diagram depicts the possibility in which you steal an apple and go to jail. The model theory argues that this initial possibility is central to the meanings of conditionals: it concerns the situation in which both the *if*-clause and the *then*-clause are true. People are faster to consider initial possibilities than alternative ones (Santamaría & Espino, 2002) and they generate more inferences from them (Johnson-Laird & Byrne, 1991; Thompson & Byrne, 2002). Moreover, if background knowledge renders the possibility incoherent, then they consider the conditional false. For example, consider these two conditionals:

1a. If Val was born in Madrid, then she was born in Spain.
b. * If Val was born in Vancouver, then she was born in Spain.

The conditional in (1a) is sensible and true – if reasoners can construct a simulation of its *if-clause,* its *then*-clause follows of necessity (Quelhas, Rasga, & Johnson-Laird, 2017). The conditional in (1b) is false, and mental models explain why: there exists no coherent scenario in which Val was born in Vancouver and Spain at the same time.

The model theory argues that when necessary, reasoners can override their default interpretation of a conditional to consider alternative possibilities, such as the possibility in which the *if*-clause is false. With some additional effort, they can flesh out their representation of *if A then B* to consider these three situations:

```
   A          B
¬  A          B
¬  A       ¬  B
```

In the diagram, each row represents a separate possibility, and the '¬' symbol represents the negated information in that clause. The only combination not represented in the diagram above is one in which *A* is true and *B* is false – and so any such situation is incompatible with the conditional. Hence, people can interpret *if you steal an apple then you go to jail* as equivalent to:

It's possible that you steal an apple and go to jail; and
it's possible that you don't steal an apple but go to jail for some other reason; and
it's possible that you don't steal an apple and don't go to jail.

Mentally simulating these additional possibilities is taxing, and so reasoners tend to consider only the initial model (in bold above).

The model theory argues that people interpret the negation of a conditional, *it's not the case that if A then B*, by constructing an initial mental model. They often reason on the basis of this initial model (Khemlani, Orenes, & Johnson-Laird, 2014). Negations of conditionals, such as *it's not the case that if you steal an apple you go to jail*, are challenging because the negation applies to the entire sentence, not to any specific clause (Khemlani et al., 2014; see also Horn, 2001). Sentential negations and their underlying sentences should be logical mirror images of each other: a sentential negation should be true whenever the sentence it negates is false, and vice versa (Horn, 2001). So how do people understand the negated conditional above? The model theory posits that they often reduce the scope of the negation, e.g., by interpreting it as meaning *if you steal an apple, then you don't go to jail.* This "small-scope" interpretation treats *it's not the case that if A then B* as: *if A then not B.* The reason for the shortcut is because it allows people to rapidly build an initial model, since the initial model of *if A then not B* is merely:

A	¬ B

This possibility is precisely what is incompatible with the affirmative conditional, *if A then B.* Reasoners who take this shortcut will treat *if A then B* and *it's not the case that if A then B* as incompatible with one another.

The small scope interpretation can lead to certain confusions and errors, because *if A then not B* is compatible with other scenarios as well, such as those in which *A* is false (Espino & Byrne, 2012; Khemlani et al., 2014). And *if A then B* is also compatible with those situations, so *if A then B* and *if A then not B* are not true logical mirror images of each other. Consider the conditionals in (2):

2a. If Val travels to Portugal, she visits Coimbra.
b. If Val travels to Portugal, she doesn't visit Coimbra.

Suppose that Val never visits Portugal. Is (2a) false and (2b) true? Or is (2a) true and (2b) false? Neither seem to be appropriate conclusions to draw. Hence, the two are not logical opposites of one another. It is also unreasonable to conclude that (2a) and (2b) are both true – perhaps a better conclusion is that both conditionals are possibly true (see Johnson-Laird, Khemlani, & Byrne, under review).

Reasoners have a more laborious way to interpret the negated conditional: they can take the complement of all possible models based on the *if*- and *then*-clauses in *if A then B* (see Khemlani et al., 2014), i.e., they can compare these models:

A	*B*
¬ *A*	*B*
¬ *A*	¬ *B*
A	¬ B

against the models of *if A then B*, which we have italicized above, to find their complement:

A	¬ B

Of course, because this enumeration process is taxing, reasoners may err in executing it.

The model theory explains the variety of ways in which people make mistakes in reasoning about negation. But, despite the inherent difficulty in processing negation (Wason, 1965), the theory makes a surprising prediction: most reasoners, most of the time, should treat *it is not the case that if A then B* as though it refers to the *A and not-B* possibility above. The two underlying processes coincide: reasoners may take an intuitive small scope shortcut, or they may reason deliberatively and enumerate all the different possibilities. The result is the same: reasoners interpret negated conditionals as referring to a single model. A direct consequence of the theory is that it predicts that *reasoners should treat conditionals as inconsistent with their negations*. They should infer that there exists no scenario compatible with both *if A then B* and *it is not the case that if A then B*.

Probabilistic theories and the model theory therefore make two opposing predictions:

- ***Probabilistic prediction:*** *Since conditionals and their negations both describe conditional probabilities, they can have non-zero probability at the same time, i.e., they can be consistent with one another..*
- ***Model theory prediction:*** *Conditionals concern coherent possibilities; they cannot be true when their negations are true and vice versa.*

We conducted two experiments to test these countervailing predictions: each study supported the model theory's account.

Experiment 1

Explicit sentential negations can confuse participants (see Khemlani et al., 2014), i.e., phrases like "it is not the case that" sound like legalese jargon in English, Spanish, and many other languages. Unlike in English, many sentences in Spanish can be negated with a preverbial *'no'*, e.g.,

3a. Soy de Cuba.	[I am from Cuba.]
b. No soy de Cuba.	[I am not from Cuba.]

But this preverbial *no* cannot be used to negate conditionals, e.g., this sentence:

* No si Alicia estudia mucho, entonces aprueba el examen.
[* Not if Alicia studies hard, she passes the exam.]

is ungrammatical. Hence, Experiment 1 placed conditional negations in the context of denials, where one individual denies the claims of another (using the Spanish *negó* = he/she denied). Denials are logically equivalent to sentential negations, but are easier to understand. The experiment gave participants problems that described the conditional claims of two different individuals, as well as a fact. Here is an example problem (translated from Spanish):

Felipe says: If Alicia studies hard, she passes the exam.
Lucas denied that if Alicia studies hard, she passes the exam.
Suppose Alicia studies hard and does not pass the exam.
In this situation, who is right?

Participants registered their response by selecting one of four possibilities:

[ ] *Felipe is right.*
[ ] *Lucas is right.* **[model theory]**
[ ] *Both are right.* **[probabilistic theory]**
[ ] *Neither of them is right.*

The model theory predicts that because the facts describing Alicia's failure correspond to the mental model of Lucas's denial, participants should infer that Lucas was right. Under the probabilistic theory, both Lucas and Felipe could be right in the event P(*passes exam* | *Alice studies hard*) > 0 and P(*does not pass exam* | *Alice studies hard*) > 0. Indeed, such accounts argue that particular outcomes can help people adjust their confidence in the conditional claims without overturning them. Hence, people should infer that both Felipe and Lucas were right.

Method

Participants. 27 undergraduates at the University of La Laguna, Tenerife, participated in the study for research credits (24 women, 3 men, mean age = 21, age range = 19-37). None of the participants had received prior instruction in logic nor had they taken part in similar experiments. This experiment and the subsequent ones received ethical approval prior to their commencement from the Committee on Ethics in Research and Animal Welfare (Comité de Ética de la Investigación y Bienestar Animal) at the University of Laguna.

Design and materials. Participants received 6 vignettes: 3 experimental vignettes designed to test their interpretations of negated conditionals, and 3 filler vignettes designed to check whether they were answering sensibly. Participants received 1 additional practice trial. The experimental problems appeared as above, and consisted of three sentences, an individual's conditional claim, another individual's denial of that claim, and a piece of evidence hypothesized to confirm the latter, in the following structure: *Person 1 says that if A then B; Person 2 denies that if A then B. A and not B happens. In this situation, who is right?* Fillers differed from experimental problems in that they described conjunctive and disjunctive claims, e.g. (translated from Spanish),

Marcos says: Alexia plants roses or plants carnations, or plants both.
Carlos says: Alexia plants roses and plants carnations.
Suppose Alexia plants roses and she plants carnations.

The task was the same: participants selected one of four possible responses to describe who was right. For the filler example above, the correct answer is that both Marcos and Carlos are right. The correct answers for the second and third filler problems were that *person 1* was right and that *neither* was right, respectively. None of the filler problems had a correct answer in which *person 2* was right.

Procedure. The problems were presented online using PsyToolkit (Stoet, 2010, 2017) in a different random order to each participant. The problems were presented sequentially, and participants had to register their response in order to move onto the next problem. The study randomized the four response options.

Open science. Experimental code, data, materials, and analysis scripts for Experiments 1 and 2 are available at https://osf.io/rkmt8/.

Results and discussion

For brevity, we describe analyses of participants' tendency to select *person 2* as right, i.e., the pattern predicted by the model theory. Participants selected *person 2* far more often for experimental than filler items (79% vs. 11%; Wilcoxon test, z = 4.38, p < .001, Cliff's δ = .89). Indeed, they selected *person 2* reliably more often than chance for experimental problems (Wilcoxon test, z = 4.55, p < .001, Cliff's δ = .93) and reliably less often than chance for filler problems (Wilcoxon test, z = 3.06, p = .002, Cliff's δ = .48). Participants exhibited the preference both in aggregate performance, as well as individually: 23 out of 27 participants selected the *person 2* response more often than any other response for experimental problems (binomial test, p < .001).

Analysis of participants' performance on filler problems revealed that they understood the task at hand: they selected the response appropriate to the particular filler item 96% of the time (Wilcoxon test against chance, z = 4.59, p < .001, Cliff's δ = .93). And 26 out of 27 participants exhibited this pattern (binomial test, p < .001).

In sum, participants responded in line with the model theory: their responses are sensible if they treated *if A then B* and its denial as inconsistent with one another. If the two statements are consistent, then they should have judged that both persons in each problem were correct. Nevertheless, one limitation of Experiment 1 is that it did not ask participants to explicitly evaluate the consistency of the conditional against its denial. Experiment 2 accordingly did so.

Experiment 2

Experiment 2 was akin to Experiment 1, except that it presented participants with a task in which they assessed the particular situations that would have to occur for the two to be consistent. Here is an example problem (translated from Spanish):

Felipe says: If Alicia trains a lot, then she takes part in the competition.
Lucas denied that if Alicia trains a lot, then she takes part in the competition.
What situation would have to occur for Felipe and Lucas to both be correct?

Alice trains a lot and she takes part in the competition.
Alice trains a lot and she doesn't take part in the competition.
Alice doesn't train a lot and she takes part in the competition.
Alice doesn't train a lot and she doesn't take part in the competition.
A situation in which both are right at the same time is not possible.

The question in each problem presupposes that a situation may exist in which the two individuals are both correct. But if participants treat a conditional and its denial as inconsistent with one another, they should select the last (bolded) option in the example above. The task above presents a stringent test of the model theory, since the it highlights every alternative option that participants could consider in which the conditional and its denial are both true.

Method

Participants. 30 participants (22 women and 8 men) different from the previous experiments participated in exchange for course credits in the University of La Laguna. Their average age was 20 years, with a range from 19 to 27 years..

Design, procedure, and materials. The design, materials, and procedure were similar to Experiment 1, but participants in Experiment 2 received 8 separate vignettes: 4 experimental problems and 4 controls. Experimental problems were similar to those used in the previous experiment: they paired a conditional with its denial. But unlike in the previous study, control problems used biconditional descriptions as well as conjunctions. Here is an example control problem (translated from Spanish):

Jacob says: If and only if there are no lions in the zoo, then there are no tigers.
Benito says: There are no lions and there no tigers at the zoo.
What situation would have to occur for Jacob and Benito to both be correct?

[ ] *There are lions and there are tigers in the zoo.*
[ ] *There are lions and there are no tigers in the zoo.*
[ ] *There are no lions and there are tigers in the zoo.*
[ ] *There are no lions and there are no tigers in the zoo.*
[ ] *A situation in which both are right at the same time is not possible.*

This problem uses explicit negations (e.g., no hay leones = "there are no lions") along with biconditionals, making it more similar in structure to the experimental problems. In the example, the correct answer (bolded) corresponds to the conjunction not-A & not-B. The four control problems varied in terms of which answer was correct, i.e., the correct answer to the first control problem was A & B, to the second it was A & not-B, and so on. As in the previous studies, the experiment randomized the order of the problems for each participant, and it presented them sequentially. The order of the five options was also randomized.

	Response options				
Problem	***A&B***	***A&¬B***	***¬A&B***	***¬A&¬B***	***Impossible.***
Experimental	3	1	8	28	**59**
Control	24	25	**27**	24	0

Table 1. The percentages of participants' responses in Experiment 2 as a function of the five separate response options and the type of problem.

Results and discussion

Table 1 shows the percentages responses as a function of the five different responses options and whether the problem was an experimental or a control problem. For brevity, we describe analyses of participants' tendency to select the *impossible* option, i.e., "A situation in which both are right at the same time is not possible."

The model theory predicts that people should select the *impossible* option for experimental but not control problems. As the table shows, they exhibited such a pattern (59% vs. 0%, Wilcoxon test, $z = 4.55$, $p < .001$, Cliff's $\delta = .76$). They selected the *impossible* option more frequently than chance (1 in 5) for experimental problems (Wilcoxon test, $z = 4.09$, $p < .001$, Cliff's $\delta = .53$) and exhibited the opposite pattern for control problems (Wilcoxon test, $z = 5.48$, $p < .001$, Cliff's $\delta = 1.0$). And 16 out of 30 participants selected the *impossible* option on 50% or more of their experimental trials (binomial test, $p < .001$). Likewise, analysis of the individual control problems shows that participants provided logically correct answers on 97% of responses, demonstrating that they understood the task.

Experiment 2 served as a strong test of the model theory's prediction: despite the fact that the experimental task was designed to discourage participants from chosing the *impossible* option, they did so for most (59%) of the experimental trials. The study controlled for several confounds: both experimental and control problems explicitly used the Spanish word *si* (= "if"), control problems used more negations than did experimental problems, and the task participants carried out asked them to selected from mutually exclusive scenarios instead of asking them to assesses who is right or evaluate the consistency of a set of assertions. Nevertheless, participants answered in a manner, both in aggregate and individually, that suggests they treat conditionals and their negations are inconsistent with one another. In sum, the results of Experiments 1 and 2 corroborate the model theory.

General discussion

Consider the following dialogue:

Amelia says: If you smoke a cigarette, you get cancer.
Elena says: No, that's not true.

What does Elena intend to communicate in her denial? Denials of simple propositions, such as, *Amelia is rich*, may be easy to comprehend, particularly in contexts in which people have information in mind to activate situations consistent with the denial (e.g., that Amelia has very little savings; see Orenes, Moxey, Scheepers, & Santamaría, 2016). Denials of compound assertions of two clauses, in particular denials of conditionals, are more difficult to understand because it is unclear to which clause the denial applies. One account argues that, because conditionals and all assertions are inherently probabilistic, they should be treated as conditional probabilities. On this account, Elena intends to communicate that the probability that you don't get cancer if you have a cigarette, P(*not-cancer* | *smoke-a-cigarette*), is high (see Handley et al., 2006; Pfeifer & Tulkki, 2017). An alternative account (Khemlani et al., 2012, 2014) argues that Elena intends to communicate what's possible, namely:

It's possible to smoke a cigarette and not get cancer.

That is, the goal of Elena's denial is to bring to light a possibility *inconsistent* with what Amelia says. Naïve reasoners, i.e., those with no training in logic or probability, have no difficulty assessing the consistency of a set of statements (Johnson-Laird, Legrenzi, & Girotto, 2004). Possibilities can conflict with one another, e.g., the situation in which Amelia is rich is inconsistent with the situation in which she has no assets. Theories based on possibilities therefore argue that conditionals should conflict with their denials. In contrast, probabilities cannot conflict with one another except at the margins (e.g., probabilities of 0.0 or 1.0). For instance, the two conditional probabilities, P(*B* | *A*) and P(*not-B* | *A*), are consistent with one another so long as they add to 1.0. Hence, theories based on probabilities predict that conditionals and their denials need not conflict.

We describe two experiments in which participants assessed that conditionals of the form, *if A then B*, were inconsistent with their denials. In each experiment, participants read vignettes describing two individuals who disagree with one another. These disagreements led participants to select one individual as right, and the other as wrong (Experiment 1). And they explicitly evaluated the individuals as being inconsistent with one another (Experiment 2).

The results reveal other, more stringent ways to test the claims of the model theory, since the theory makes distinct predictions about how people reason with negations and conditionals. That is, the model theory posits that the interpretation of negations depends heavily on the initial set of mental models that reasoners construct. Recent chronometric evidence supports this claim: Orenes, Moreno-Ríos, and Espino (2022) show that people read the conjunction *A & B* just as quickly after reading *if A then B* compared to when they read *it is not the case that A and B*, which suggests that they represent some trace of the *A & B* possibility even when it is negated. If the initial possibilities they construct bias how they reason, then it can affect how they interpret the negations of conditional expressions – different sets of initial mental models will produce different patterns of reasoning. Many factors can affect the initial set of models people construct (Johnson-Laird & Byrne, 2002), including the contents of the terms in the expression (Orenes & Johnson-Laird, 2012; Quelhas, Johnson-Laird, & Juhos, 2010); temporal and spatial relations between antecedent and consequence (Juhos, Quelhas, & Johnson-Laird, 2012); and any disabling conditions implied by the contents (Gómez-Sánchez, Moreno-Ríos, Couto, & Quelhas, 2021).

Perhaps more recent probabilistic accounts of the conjunction of conditionals can address the dilemma outlined above. Sanfilippo et al. (2020) assume that the conditionals *if A then B* and *if A then not B* are best modeled by P(*B* | *A*) and P(*not-B* | *A*), respectively. Over (personal communication) observes that under this account, the probability of a conditional conjunction, e.g., P(*if A then B* & *if A then not-B*) is always 0, and hence conditionals and their negations are inconsistent with one another. Indeed, Sanfilippo et al. (2020, p. 170) provides this formula to compute the probability of conditional conjunctions:

$$P[(A|H)\&(B|K)] = P[(A|H) * (B|K)] = P(A|H)*P(B|K)$$

Yet, if we replace ($A|H$) with ($B|A$) and ($B|K$) with (*not-B* | *A*), the equation above resolves to:

$$P(B|A) * P(not\text{-}B|A)$$

In the case that both conditional probabilities are .50, then the value of the above is .25, i.e., above zero; by Sanfilippo et al.'s own argument, *if A then B* and *if A then not B* are consistent with one another.

Given the frequency with which people interpret negations of conditionals, *it is not the case that if A then B*, as *if A then not-B*, readers may wonder whether this interpretation suffices as a sensible, and even normative, treatment, as researchers have suggested (Handley, Evans, & Thompson, 2006; Pfeifer & Tulkki, 2017). If so, then theories that treat such interpretations as normative must explain how and why people deviate from such norms. In contrast, the model theory espouses a different view: *if A then not-B* is a reasonable heuristic interpretation of the negated conditional. The heuristic is effective because it can help people rapidly construct relevant mental models – but, as we outline in the Introduction, it can lead to systematic mistakes. The model theory therefore predicts that people's interpretations of negated conditionals should be varied (see Khemlani et al., 2012, 2014) and they should be traceable, i.e., they should affect nonverbal behaviors, such as the time it takes people to hear passages and the time it takes them to process visual imagery (see Orenes, 2021).

In sum, while denials of conditionals may pose difficulties in interpretation, they are not rare in daily speech. Reasoners routinely deny conditional claims, and they interpret the meanings of such denials systematically. In this paper, we showed evidence of this systematicity: people treat denials of conditionals and the conditionals they deny as inconsistent from one another.

Acknowledgments

This research was supported by funding to Moreno-Ríos from the Spanish Government, Ministry of Economy and Competitiveness (MINECO) Grant PGC2018-095868-B-I00), to Orenes from the Spanish Government (MINECO Grant RTI2018-098730-B-I00), and to Khemlani from the Naval Research Laboratory to study epistemic reasoning. For their comments and criticisms on earlier versions of this work, we thank Bill Adams, David Beltrán, Branden Bio, Gordon Briggs, Ruth Byrne, Hillary Harner, Tony Harrison, Phil Johnson-Laird, Laura Kelly, Markus Knauff, Andrew Lovett, Marco Ragni, and Greg Trafton.

References

Baratgin, J., Douven, I., Evans, J. St. B. T., Oaksford, M., Over, D., Politzer, G., & Thompson, V. (2015). The new paradigm and mental models. *Trends in Cognitive Sciences*, 19.

Byrne, R. M. J., & Espino, O. (2021). Conditional and counterfactual reasoning. In M. Knauff and W. Spohn (Eds.), *The Handbook of Rationality*. Cambridge, MA: MIT Press.

Byrne, R. M. J., & Johnson-Laird, P. N. (2020). If and or: Real and counterfactual possibilities in their truth and probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46.

Cooper, W. S. (1968). The propositional logic of ordinary discourse. *Inquiry*, 11, 295-320.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, 101, 50–81.

Espino, O. & Byrne, R. (2012). It is not the case that if you understand a conditional you know how to negate it. *Journal of Cognitive Psychology*.

Espino, O., Santamaría, C., & Byrne, R. M. J. (2009). People think about what is true for conditionals, not what is false: Only true possibilities prime the comprehension of 'if'. *Quarterly Journal of Experimental Psychology*, 62, 1072-1078.

Evans, J.S.B.T., Handley, S.J., & Over, D.E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 321–335.

Evans, J. S. B., & Over, D. E. (2004). *If: Supposition, pragmatics, and dual processes.* Oxford University Press, USA.

Evans, J. S. B., & Over, D. E. (2013). Reasoning to and from belief: Deduction and induction are still distinct. *Thinking & Reasoning*, 19.

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition and logical form.* Academic Press: New York.

Gómez-Sánchez, J., Moreno-Ríos, S., Couto, M., & Quelhas, A.C. (2021). Conditional content, explicit information and generating cases: Sources for suppressing inferences. *Acta Psychologica*, 213, 103420.

Goodwin, G. P., & Johnson-Laird, P.N. (2018). The truth of conditional assertions. *Cognitive Science*, 42, 2502–2533.

Grice, P. (1989*). Studies in the Way of Words.* Harvard University Press.

Handley, S. J., Evans, J. St. B. T., & Thompson, V. A. (2006). The negated conditional: A litmus test for the suppositional conditional? *Journal of Experimental Psychology: Learning, Memory, Cognition*, 32, 559–569.

Horn, L. R. (2001). *A natural history of negation*. Stanford, CA: CSLI.

Jeffrey, R. (1981*). Formal logic: Its scope and limits (2nd ed.).* New York: McGraw-Hill.

Johnson-Laird, P. N., & Byrne, R. M. (1991). *Deduction*. Lawrence Erlbaum Associates, Inc.

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

Johnson-Laird, P. N., Legrenzi, P., & Girotto, V. (2004). How we detect logical inconsistencies. *Current Directions in Psychological Science*, 13.

Johnson-Laird, P.N., Khemlani, S., & Byrne, R.M.J. (under review). Truth, verification, and reasoning. Manuscript under review.

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, 19.

Juhos, C., Quelhas, A. C., & Johnson-Laird, P. N. (2012). Temporal and spatial relations in sentential reasoning. *Cognition*, 122, 393-404.

Khemlani, S. S., Byrne, R. M., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science*, 42, 1887-1924.

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2012). Negation: A theory of its meaning, representation, and use. *Journal of Cognitive Psychology*.

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2014). The negation of conjunctions, conditionals, and disjunctions. *Acta Psychologica*, 151.

Kleiter, G., Fugard, A., & Pfeifer, N. (2018). A process model of the understanding of conditionals. *Thinking & Reasoning*, 24, 386-422.

López-Astorga, M., Ragni, M., & Johnson-Laird, P. N. (2021). The probability of conditionals: A review. *Psychonomic Bulletin & Review*.

Nickerson, R. S. (2015). *Conditional reasoning: The syntactics, semantics, thematics, and pragmatics of 'If'.* New York: Oxford University Press.

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

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29.

Orenes, I. (2021). "Looking at" negation: Faster processing for symbolic rather than iconic representations. *Journal of Psycholingustic Research*, 50, 1417-1436.

Orenes, I., & Johnson –Laird, P.N. (2012). Logic, models, and paradoxical inferences. *Mind and Language*, 27, 357-377.

Orenes, I., Moreno-Ríos, S., & Espino, O. (2022). Representing negated statements: When false possibilities also play in the mind. *Journal of Cognitive Psychology*, 34, 1052-1062.

Orenes, I., Moxey, L., Scheepers, C., & Santamaría, C. (2016). Negation in context: Evidence from the visual world paradigm. *Quarterly Journal of Experimental Psychology*, 69, 1082-1092.

Partee, B. H. (1979). Semantics—mathematics or psychology?. In *Semantics from different points of view* (pp. 1-14). Springer, Berlin, Heidelberg.

Pfeifer, N., & Tulkki, L. (2017). Conditionals, counterfactuals, and rational reasoning: An experimental study on basic principles. *Minds and Machines*, 27, 119-165.

Quelhas, A. C., Johnson-Laird, P.N., & Juhos, C. (2010). The modulation of conditional assertions and its effects on reasoning. *Quarterly Journal of Experimental Psychology*, 63, 1716-1739.

Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2017). A priori true and false conditionals. *Cognitive Science*, 41, 1003-1030.

Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2018). The relation between factual and counterfactual conditionals. *Cognitive Science*, 42, 2205-2228.

Ramsey, F. P. (1929/1990). General propositions and causality (originally published 1929). In D. H. Mellor (Ed.), *Philosophical papers* (pp. 145-163). Cambridge: Cambridge University Press.

Sanfilippo, G., Gilio, A., Over, D. E., & Pfeifer, N. (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning*, 121, 150-173.

Santamaría, C., & Espino, O. (2002). Conditionals and directionality: On the meaning of if vs. only if. *Quarterly Journal of Experimental Psychology Section A*, 55, 41-57.

Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42.

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*.

Thompson, V. A., & Byrne, R. M. (2002). Reasoning counterfactually: making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 1154.

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 4.