COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2017) 1–28

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12479

A Priori True and False Conditionals

Ana Cristina Quelhas,[a] Célia Rasga,[a] Philip N. Johnson-Laird[b,c]

[a]*William James Center for Research, ISPA-Instituto Universitário*
[b]*Department of Psychology, Princeton University*
[c]*Department of Psychology, New York University*

Received 23 May 2016; received in revised form 31 October 2016; accepted 21 November 2016

Abstract

The theory of mental models postulates that meaning and knowledge can modulate the interpretation of conditionals. The theory's computer implementation implied that certain conditionals should be true or false without the need for evidence. Three experiments corroborated this prediction. In Experiment 1, nearly 500 participants evaluated 24 conditionals as true or false, and they justified their judgments by completing sentences of the form, *It is impossible that A and* ___ appropriately. In Experiment 2, participants evaluated 16 conditionals and provided their own justifications, which tended to be explanations rather than logical justifications. In Experiment 3, the participants also evaluated as possible or impossible each of the four cases in the partitions of 16 conditionals: *A and C*, *A and not-C*, *not-A and C*, *not-A and not-C*. These evaluations corroborated the model theory. We consider the implications of these results for theories of reasoning based on logic, probabilistic logic, and suppositions.

Keywords: Deductive reasoning; Conditionals; Logic; Mental models; Modulation; Possibilities

1. Introduction

The meanings of "if," "and," and "or" in daily life differ from the meanings of their logical analogs. Their analogs in logic have constant meanings from one sentence to another. They map the truth values of the clauses that they connect onto a truth value for the sentence as a whole (see, e.g., Jeffrey, 1981). For instance, the analog of *if-then* in logic is material implication, and a sentence such as:

It's hot materially implies it's raining.

Correspondence should be sent to Ana Cristina Quelhas, ISPA-Instituto Universitário, William James Center for Research, Rua Jardim do Tabaco, 34, 1149-041 Lisboa, Portugal. E-mail: cquelhas@ispa.pt

is true in any case except one in which it is hot but not raining. In daily life, the meanings of connectives in general, and conditionals in particular, do not refer to truth values (Johnson-Laird & Byrne, 1991). And their meanings are not constant, but vary from one assertion to another (Nickerson, 2015). What happens according to the theory of mental models is that knowledge and the meanings of clauses in sentences *modulate* the interpretation of *if* and other connectives (Johnson-Laird & Byrne, 2002). We explain presently how this process works, at least as it is embodied in a computer program implementing the theory, and we emphasize that modulation has been corroborated in experiments (Goodwin & Johnson-Laird, 2005; Johnson-Laird & Byrne, 2002; Juhos, Quelhas, & Johnson-Laird, 2012; Quelhas & Johnson-Laird, 2016; Quelhas, Johnson-Laird, & Juhos, 2010). The aim of this study was to show that it has a still more radical effect that has hitherto been overlooked. It can establish the truth or falsity of compound assertions a priori, that is, they are true or false independently of any empirical evidence. All of us, for example, are likely to judge that this conditional is true:

If Pat is reading the article, then Pat is alive.

Conditionals such as the preceding one are *specific* because they refer to particular individuals, whereas conditionals such as:

If a person is reading an article, then the person is alive.

are general. The crucial difference between these two sorts of conditional is elucidated by the *partition* for a conditional, *if A then C*, such as:

A	C	Pat is reading this article and Pat is alive.
A	not-C:	Pat is reading this article and Pat is not alive.
not-A	C:	Pat is not reading this article and Pat is alive.
not-A	not-C:	Pat is not reading this article and Pat is not alive.

Each case in the partition is a conjunction, such as *A & C*, but for simplicity we omit the sign for conjunction. The second of the preceding cases is impossible, but the facts of the matter for a specific conditional must be just one of the remaining cases in the partition. In contrast, the facts of the matter for a general conditional can correspond to more than one case in the partition; for the general conditional above, there can be persons reading the article who are alive, persons who are not reading the article who are alive, and persons who are not reading this paper who are not alive. As a consequence, the logic of general conditionals is more powerful than the logic of specific conditionals (Jeffrey, 1981). As readers will see, the concept of a partition plays an important role in the present paper.

The rest of this introduction describes the controversy in philosophy about a priori truths. It then outlines the unified theory of mental models, which unifies the role of models in reasoning about facts, possibilities, and probabilities. It explains the theory's mechanism for modulation, which is implemented in a computer program for reasoning with sentential connectives, *mSentential*, for which the source code can be downloaded from http://mentalmodels.princeton.edu. It describes the theory's predictions about descriptive versus causal conditionals, and about specific versus general conditionals. It reports three experiments examining the evaluations of conditional assertions, which show that the process does produce a priori truth values. These results also corroborate the unified model theory's account of the meanings of conditionals as referring to conjunctions of possibilities. Finally, the paper considers the implications of the results for other theories of human reasoning, including those based on logic (see, e.g., Braine & O'Brien, 1998; and Rips, 1994), on suppositions (e.g., Evans, 2007), and on probabilistic logic (see, e.g., Cruz, Baratgin, Oaksford, & Over, 2015; Evans, 2012; Pfeifer & Kleiter, 2009).

1.1. The philosophical controversy about a priori truth values

The idea that certain assertions have a priori truth values has a long history in philosophy, though philosophers have focused on truths more than falsehoods. In the 17th century, Leibniz (1686/2002) distinguished between logical truths and contingent truths (aka *synthetic* truths). A logical truth, such as:

It is raining or it is not raining

is bound to be true granted the meanings of *not* and *or* in logic. A contingent assertion, such as:

It is raining or it is hot

is true or false depending on the state of the world. In the 18th century, Hume (1739/1978) drew a similar distinction. And, most famously, Kant (1781/1934) accepted the distinction, but distinguished between two sorts of a priori truths. *Analytic* truths are akin to logical truths, but they are true in virtue of meanings, for example:

A triangle has three sides.

Synthetic assertions are true a priori depending on the nature of the world and include such examples as:

All bodies are extended in space.

Kant also included in this category the propositions of Euclidean geometry—non-Euclidean geometry was unknown to him. In the 20th century, philosophers continued to defend a priori truths (e.g., Carnap, 1947) until Quine (1951) published an influential critique (republished in Quine, 1953). He argued that the distinction between a priori and contingent truths was an unempirical dogma as was the notion that propositions could be reduced to constituent parts. He allowed that some propositions are true in virtue of logic, for example:

No unmarried man is married.

But he took exception to the idea that:

No bachelor is married

is true a priori. To the argument that "bachelor" is synonymous with "unmarried man," he retorted that the notion of synonymy was just as much in need of clarification as analyticity itself. Since the era in which he was writing, however, linguists and psychologists have made some progress in establishing the required notion of cognitive synonymy (see, e.g., Chomsky, 1977; Miller & Johnson-Laird, 1976).

Quine argued further that one cannot assess the truth or falsity of isolated propositions: "The unit of empirical significance is the whole of science" (Quine, 1953, p. 42). No statement, such as the one above about bachelors, has a truth value that is immune to revision. And no principled distinction exists between analytic and synthetic assertions: "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (Quine, 1953, p. 41). Quine's skepticism elicited a large philosophical literature. In daily life, however, people do seem to take for granted that an assertion such as "The waiter stole the client's wallet" can be judged true or false by itself. They also appear to treat some assertions as true solely as a result of their meanings. This ability is critical in many situations—from the assessment of evidence in legal proceedings to arguments in everyday life. But, philosophers who think Quine is right—of which there are many—might well argue that people are mistaken, and unaware of the true complexities of meanings.

Our intention is not to grapple with philosophical niceties, but rather to show that the distinction between assertions that have a truth value a priori and those that do not is no longer an unempirical dogma, and to explain how its existence is a consequence of modulation, at least for naive individuals. And by "naive," we mean merely individuals who have not studied philosophy or logic, and whose judgments therefore reflect common sense about everyday life.

1.2. The unified theory of mental models

The unified theory of mental models postulates that assertions refer to possibilities, which are represented in mental models, and that reasoners use these models to make inferences (see, e.g., Johnson-Laird, 1983, 2006; Johnson-Laird & Byrne, 1991; Khemlani & Johnson-Laird, 2013). The theory distinguishes between *mental* models, which represent only what is true given the premises, and *fully explicit* models, which also represent what is false. The conditional:

If the triangle is present, then the circle is present

yields a conjunction of two mental models, depicted here on separate rows:

△ ○
. . .

The first and principal model represents the possibility in which the triangle and the circle are both present. The second model, which the ellipsis denotes, has no explicit content and represents the possibilities in which the *if*-clause of the conditional, *the triangle is present*, is false. Mental models underlie intuitions. But the fully explicit models of the conditional, which deliberations can construct, are as follows, in the order in which they appear to be available to individuals (e.g., Barrouillet, Grosset, & Lecas, 2000; Johnson-Laird & Byrne, 2002):

△	○
not-△	not-○
not-△	○

where “not” stands for negation, for example, “the triangle is not present.” These representations have a structure isomorphic to the structure of the situations to which the assertions refer; that is, the conditional refers to a conjunction of three distinct possibilities, and each fully explicit model corresponds to one of them (Hinterecker, Knauff, & Johnson-Laird, 2016). Of course, mental models are models of the world, not diagrams such as the preceding ones. Models specify as little as possible. The first model depicted above, for example, is compatible with infinitely many alternatives in which the triangle (of unknown shape, size, color, etc.) is present with the circle (also of unknown shape, size, color, etc.).

The evidence for the unified theory is robust (see, e.g., Johnson-Laird, Khemlani, & Goodwin, 2015), but for our purposes the critical feature of the theory is modulation. It is the process by which knowledge of meanings, context, and the world, which itself is represented in models, can block the construction of models of assertions or add information to them. Consider, for instance, the difference between these two contingent assertions:

If it is raining, then it is hot

and:

If it raining, then it is pouring.

Any of the three possibilities to which the first conditional refers could be the case, but the second conditional refers only to two possibilities. To see why, consider the three possibilities analogous to those for the first conditional:

raining pouring
not-raining not-pouring
not-raining pouring

The third case is, in fact, impossible, because the meaning of “pouring” implies that it is raining. This meaning yields the following models of possibilities in knowledge:

pouring	raining
not-pouring	not-raining
not-pouring	raining

Modulation uses these models in knowledge to interpret assertions. They block the construction of a model in which it is pouring but not raining. The result of modulation is accordingly equivalent to the intersection of the two sets of models, one set for the assertion and the other set in knowledge. The two sets in our example have in common only two models:

raining	pouring
not-raining	not-pouring

They correspond to a biconditional interpretation, *if and only if it's raining, then it's pouring*.

Empirical studies have shown that modulation yields many different interpretations for conditionals (see Goodwin & Johnson-Laird, 2005; Johnson-Laird & Byrne, 2002; Juhos et al., 2012; Quelhas et al., 2010) and for disjunctions (Quelhas & Johnson-Laird, 2017). But critics often suggest that knowledge could contain, not models, but representations of propositions akin to sentences in a belief box, for example:

If it is pouring, then it is raining.

The problems with this assumption are two-fold. First, the sentences in the box have to be understood, that is, to yield a representation common to them and to synonymous assertions, for example, *it's raining, if it's pouring*. Second, some sorts of modulation depend on models that would have to be envisaged before they could be formulated as verbal descriptions. For example, experiments have shown that modulation can add temporal and other relational information to models. A conditional, such as:

If Lisa received the money, then she paid Frederico.

elicits the interpretation that Lisa's possible receipt of the money occurred before her possible payment to Frederico. In contrast, the conditional:

If Mauro did well on the exams, then he studied a lot.

yields an interpretation in which Mauro's possible performance in the exams occurred after—and presumably because—he studied a lot (Quelhas et al., 2010). Modulation can also introduce spatial relations, and for a conditional, such as:

If the maid cleans the desk, then she moves the folders to the floor

it yields both a temporal and spatial relation in a kinematic simulation: The maid first moves the folders from the desk to the floor, and then she cleans the desk (Juhos et al., 2012). To suppose that such modulations depend on verbal propositions is to reverse the proper order of psychological processes. To formulate the appropriate verbal propositions, one would first have to simulate the sequence of events. But it is precisely this simulation that the preceding conditional elicits to interpret the conditional. There is normally no need to go on to formulate a verbal description of the simulation. Modulation is therefore a process that depends on knowledge-based models, not propositions, where knowledge is about meanings and about the world.

The burden of the previous account is that models in knowledge can modulate the interpretation of a conditional, such as *If it is raining, then it is pouring*, so that it is a contingent biconditional. But one can also ask about the effect of models in knowledge on the converse conditional:

If it is pouring then it is raining.

This conditional refers to exactly the set of possibilities that are already in knowledge, which derive from the meanings of "pouring" and "raining." It follows that the conditional is true: Its meaning guarantees its truth a priori. We noticed this coincidence in developing the *mSentential* computer program, which implements reasoning and modulation, and the observation prompted the present research.An *atomic* assertion is one that is affirmative and that contains no sentential connectives, for example:

The tulip is a flower.

This assertion is true a priori. Another atomic assertion is:

The infant is an adult.

It is false a priori. These claims are not unempirical, because Steinberg (1970, 1975) showed that participants sort atomic assertions into categories that reflect their a priori truth values. They also treat assertions, such as *the moon is a newspaper* as false, and their negations as true (Steinberg, 1972). As far as we know, however, neither Steinberg nor anyone else has examined compound assertions such as conditionals to determine whether they too can have a priori truth values.

1.3. The truth and falsity of conditionals

In what circumstances is a specific conditional, *If A, then C*, true a priori and in what circumstances is it false a priori? In logic, material implication—the sentential connective closest in meaning to a conditional—is true in any case except one in which *A* is true but *C* is false (Jeffrey, 1981). Hence, *A materially implies C* is true given that *not-A* is true; and it is also true given that *C* is true. In either case, the material implication is bound to be true. You know, for example, that the following assertion is true:

Hitler is not alive.

And so if a specific conditional corresponds to material implication, it follows that:

If Hitler is alive, then the Allies lost World War II.

Most people reject this sort of bizarre inference (Orenes & Johnson-Laird, 2012). It is known as a "paradox" of material implication. For many students of reasoning, the paradoxes rule out material implication as the meaning of conditionals (e.g., Evans & Over, 2004; Johnson-Laird & Byrne, 2002; Oaksford & Chater, 2007).

The unified theory differs from logic. A specific conditional, *If A, then C*, on which modulation has had no effect, refers to a conjunction of three cases as possible:

A	C
not-A	not-C
not-A	C

Modulation can eliminate any case, but if the conditional refers to at least two possibilities, one of them must correspond to its principal mental model, *A C* (Johnson-Laird & Byrne, 2002, p. 660). It follows that a conditional is true a priori provided that the cases to which it refers are all real possibilities. On this account, the conditional about Hitler is not true a priori: It is impossible that Hitler is alive and that the Allies lost the war. Hence, the case corresponding to the mental model of the conditional is impossible. In contrast, as we have already shown, the following conditional should be evaluated as true a priori:

If it is pouring, then it is raining.

Knowledge establishes the following evaluations of the cases in the partition:

It is pouring.	It is raining:	possible
It is pouring.	It is not raining:	not possible
It is not pouring.	It is raining:	possible
It is not pouring.	It is not raining:	possible

Similarly, the conditional:

If it is pouring, then it is not raining

is false a priori according to the unified theory, because the case corresponding to its mental model (It is pouring and it is not raining) is impossible.

There are also conditionals for which modulation yields a true biconditional interpretation, for example:

If Emma is illiterate, then she is unable to read.

It refers to two possibilities about Emma. She is:

Illiterate	unable to read
not illiterate	not unable to read

And there are conditionals for which modulation yields a false biconditional interpretation, for example:

If Emma is illiterate, then she is able to read.

1.4. Descriptive versus causal conditionals, and specific versus general conditionals

Descriptive conditionals, such as:

If Francis is swimming in the river, then he is in the water

can be true (or false) in virtue of the meanings of words, that is, a river is a stream of water. Of course, Quine is right that the meanings of words, such as "river," may change, but judgments of truth and falsity depend on current knowledge, and so the preceding conditional should elicit a judgment of "true." In contrast, causal conditionals, such as:

If it is raining, then the streets are wet concern knowledge of the world. They should therefore be more vulnerable to counterexamples. For instance, the preceding causal assertion would be false if the streets were under a roof, or the streets are so hot that the rain immediately evaporates. A corollary is that a priori truth values are more likely to be assigned to descriptive conditionals than to causal conditionals.

A similar difference should occur between specific conditionals and general conditionals. With a specific conditional, such as:

If Fred is a bachelor, then he is unmarried

it is natural to accept the relevant meaning of "bachelor," and so hard to envisage a counterexample. But with a general conditional, such as:

If anyone is a bachelor, then he is unmarried

the sorts of counterexamples that Quine (1953, p. 28) envisaged are more likely to come to mind, provoked by the universal claim. Hence, a person who is a bachelor of arts could be married. So a priori truth values should be more likely to be assigned to specific conditionals than to general conditionals. Our studies examined these various sorts of conditional in part to ensure that they used representative samples of different sorts of conditional, and in part to test these subsidiary predictions. The main aim, however, was to discover whether naive individuals judge that conditionals can be true or false a priori.

2. Experiment 1: A preliminary study of judgments of a priori truth values

Our first experiment tested whether naive individuals judged the truth values of conditionals according to the unified theory. The study was preliminary to determine whether or not Quine (1953) was right that no substantial grounds exist for such judgments. Hence, the experiment tested a large sample of participants—just under 500, using an on-line website, and it included causal and descriptive conditionals. After the participants had made their evaluation of each assertion, they provided a justification for it by completing a sentence of the sort:

It is impossible that A and ____________.

where *A* was the *if*-clause of the conditional. The theory predicts that participants should complete the sentence with *not C* for those assertions that they judge to be true, and with *C* for those assertions that they judge to be false.

2.1. Method

2.1.1. Participants

The experiment tested 494 participants (248 males and 246 females with mean age 45.7 years) on an online platform. All the participants worked for the same large company in Portugal, and they had different levels of education, ranging from no further education after middle school (2%), high school (50%), an undergraduate degree (41%), or a master's degree (7%). Their participation was voluntary and solicited by an email inviting them to take part in the experiment during the following 5 days via a link to the experiment on a website.

2.1.2. Design

The participants acted as their own controls and carried out the task for six instances of each of four sorts of assertion according to the unified theory: conditionals that they should evaluate as true, and those that they should evaluate as false; and conditionals that they should interpret as biconditionals and evaluate as true, and those that they should evaluate as false. Half the conditionals in the experiment were descriptive and half were causal. The 24 assertions were presented to each participant in a different random order.

2.1.3. Materials and procedure

We devised 48 assertions, so that according to the unified theory, 24 should be interpreted as conditionals and 24 should be interpreted as biconditionals, and half of them should be judged to be true a priori and half of them should be judge to be false a priori. The contents were everyday topics, and half the assertions were descriptive, such as:

If Mary has the flu, then she is sick

and half were causal, such as:

If Sarah turns off the light in the room, then it will be darker.

The predicates in the *then*-clause were usually binary to enable us to construct corresponding assertions that were false:

If Mary has the flu, then she is healthy

and:

If Sarah turns off the light in the room, then it will be lighter.

The false assertions were constructed from the true ones, by substituting the other member of the binary predicate in place of the original. We used the resulting 48 conditionals to construct two complementary sets of materials (sets I and II) in order to avoid the same participant judging both the true and the false members of a matched pair. The complete set of materials is in Table S1 of the Supporting Information file. The participants were randomly assigned to work with one of the two sets of materials.

The task was presented on the online platform Qualtrics Survey Software, and each of the 24 problems appeared on a separate screen, and each participant received them in a different random order. The key instructions were as follows (translated from the original Portuguese):

> You will be presented with 24 sentences and your task is to classify each of them as true or false. Afterwards you'll have to justify your answer by completing the sentence that will be presented to you in the available rectangle.

On each trial, the participants saw the following sort of display (again, as in all our experiments, translated from the original Portuguese):

Please consider the following sentence:

If Sara turns off the light in the room, then it will be lighter.

Do you think that the sentence is: True ☐ False ☐

Please justify your answer by completing the following sentence:

It is impossible that Sara turns off the light in the room and ______________________

The task took the participants about a quarter of an hour to complete.

2.2. Results

There was no reliable difference between the two groups used to counterbalance the materials in the percentages of their predicted judgments of truth values of the conditional assertions (97.6% and 98.0%, Mann–Whitney test, $z = 1.29$, $p > .15$). Likewise, there was no difference in the predicted judgments between the causal and descriptive conditionals, which were both at ceiling (98.0% and 98.0%, Wilcoxon test, z = 1.24, $p > .21$). Hence, we amalgamated the results for analysis. Table 1 presents the resulting frequencies and corresponding percentages of judgments of the truth or falsity of the conditionals and of the biconditionals. For all our studies, we used nonparametric statistical tests, such as the Mann–Whitney and Wilcoxon, because they call for no assumptions about the distribution of data, and because for a given number of participants they are less powerful than parametric tests, such as analysis of variance, and therefore less likely to yield a Type II error, that is, an incorrect rejection of the null hypothesis (see Siegel & Castellan, 1988).

The results corroborated the unified theory's predictions at ceiling. Wilcoxon tests corroborated the predictions overall ($z = 42$, $p <$ a one in a billion), and for each of the eight sorts of conditional, whether they were causal or descriptive (with $z > 20$ in all conditions).

Table 1
The frequencies (and percentages) of the judgments of truth values in Experiment 1 of conditionals predicted to be evaluated as true or false and to have a conditional or biconditional interpretation

		Participants' Judgments	
Predicted Judgment	Predicted Interpretation	True	False
True	Conditional	**2895 (98%)**	69 (2%)
	Biconditional	**2903 (98%)**	61 (2%)
False	Conditional	67 (2%)	**2897 (98%)**
	Biconditional	67 (3%)	**2897 (97%)**

Note. Predicted evaluations are shown in bold.

Table 2
The frequencies (and percentages) for completions of It's impossible that A and __ for conditional assertions in Experiment 1, where A stands for the if-clauses of conditionals

		Completions of It is Impossible that A and ____	
Predicted Judgment	Predicted Interpretation	Not-C	C
True	Conditional	**2834 (96%)**	49 (3%)
	Biconditional	**2842 (96%)**	91 (3%)
False	Conditional	62 (2%)	**2857 (96%)**
	Biconditional	37 (1%)	**2894 (98%)**

Note. Predicted answers shown in bold, and the balances of the percentages (1–2%) were various sorts of miscellaneous responses.

The participants' completions of the sentences to justify their judgments were also at ceiling (Table 2). Two independent raters evaluated the completions, and their concurrence was highly reliable ($k = 0.905$, $p < .001$). Because of the power of the experiment, there was a significant difference between the two groups for counterbalancing the materials (Mann–Whitney test, $z = 2.49$, $p < .025$), but the differences in each sort of conditional were never more than 2%, and so we have pooled the results for further statistical analysis. The results corroborated the unified theory's predictions again at ceiling. Wilcoxon tests corroborated the predictions overall ($z = 19.96$, $p < .0001$), and for each of the eight sorts of conditional, whether they were causal or descriptive (with $z > 14.0$ in all conditions). However, the completion *It is impossible that A and not-C* is consistent with both a conditional and a biconditional interpretation, and so we have no evidence to show that the participants distinguished between their meanings—other studies of modulation, however, have shown that individuals do distinguish the possibilities to which they refer (see, e.g., Johnson-Laird & Byrne, 2002), and, as a result, draw different inferences from them (Quelhas et al., 2010).

In sum, nearly 500 participants judged that some conditionals are true a priori and some conditionals are false a priori. Their judgments tended to corroborate the particular predictions of the unified theory at a level greater than 95%. And they tended to justify judgments of the truth of *If A, then C* by responding that *A* and *not-C* was impossible, and to justify judgments of the falsity of *If A, then C* that *A* and *C* was impossible. However, a potential weakness of the study was that the participants were highly constrained in the judgments of truth values that they could make (either true or else false) and in the justifications they could make, so how would they perform without such tight constraints? Our next experiment answered this question.

3. Experiment 2: A replication with less constraint on judgments and justifications

This experiment aimed to replicate the judgments of truth and falsity in a conventional experiment instead of an on-line one, to examine the judgments of truth values when the participants were also allowed to respond, "impossible to say," and to reveal the nature of their justifications for their judgments when they were allowed to respond in their own words.

3.1. Method

3.1.1. Participants

The participants were 28 psychology undergraduates from ISPA-IU, in Lisbon, who volunteered to take part in the experiment. They were 20 women and 8 men, average age 20.5 years (*SD* = 6.4).

3.1.2. Design

The participants acted as their own controls in judging whether the truth values of conditional assertions were true, false, or impossible to determine. According to the unified theory, half the conditionals should be judged to be true and half the conditionals should be judged to be false. The experiment crossed this manipulation with whether the conditionals were specific or general, and with whether they were descriptive or causal. The three manipulations yielded a total of eight trials for each participant. After they had judged the truth value for a conditional, they wrote a justification for their judgment.

3.1.3. Materials and procedure

We devised a set of 16 everyday conditional assertions of which, according to the unified theory, half were true and half were false. Likewise, within the set, half were specific assertions, such as:

If Ana puts a kettle over the fire, then the water inside heats up

and half were general assertions, such as:

If a person is swimming in the river, then he or she is in the water.

And within each set, half were descriptive as in the second of these examples, and half were causal as in the first of these examples. Like Experiment 1, the conditionals had binary predicates, such as *healthy* and *sick* to yield matching true and false conditionals. Table S2 in the Supporting Information file presents the full set of assertions in English and in their original Portuguese. We constructed two sets of eight assertions, and randomly assigned participants to one set or the other to ensure that they saw only one member of a matched pair of true and false conditionals. We randomly assigned participants to these two groups but ensured that there were equal numbers in each group.

The participants were tested together in a single room, and the experiment was presented to them in a booklet, with each conditional on a different page, and the eight pages in a different random order for each participant. A typical problem was as follows:

Please consider the following sentence:

If Ana puts a kettle over the fire, then the water inside heats up.

Do you consider that this sentence is: True ☐ False ☐ Impossible to say ☐

Please justify your answer between the lines below:

___

___

There was no time limit to complete the experiment. The first page asked for the age and gender of the participant and had the following instruction (translated from the Portuguese):

> With the growing interest in the study of criminology, the need to assess the reasoning capabilities of future criminologists has become important. It is vital for any criminologist to be able to analyse information and to extract meaning from them in order to draw conclusions. This task assesses your perspicacity in those abilities.

The reference to criminology aimed to motivate undergraduate participants to think carefully. The experimenter gave a short debriefing after the final participant had completed the task.

3.2. Results and discussion

The two groups for counterbalancing the materials did not differ reliably in the percentages of their predicted judgments of truth and falsity (72% vs. 69%, Mann–Whitney test, $z = 0.54$, $p > .6$). We therefore amalgamated their results for analysis. Table 3 presents the percentages of judgments of the conditional assertions. Overall, the results corroborated the unified theory's predicted judgments (71% predicted judgments vs. 29% unpredicted judgments, Wilcoxon test, $z = 3.97$, $p < .00005$, one tail). There was no reliable difference in the corroboration of conditionals predicted to be true (68%) and those predicted to be false (73%; Wilcoxon test, $z = 1.29$; $p = .198$). But, as the unified theory

Table 3
The percentages of the judgments of truth values of the eight sorts of conditionals in Experiment 2. The conditionals were predicted to be true or false, and they were general or specific, and causal or descriptive

		Participants' Judgments		
Predicted Judgment	Conditional	True	False	Impossible to Say
True	General causal	**39**	11	50
	General descriptive	**82**	7	11
	Specific causal	**68**	4	29
	Specific descriptive	**82**	0	18
False	General causal	32	**54**	14
	General descriptive	0	**82**	18
	Specific causal	0	**85**	14
	Specific descriptive	0	**71**	29

Note. Results for the predicted judgments are shown in bold.

predicts, there were more predicted judgments for descriptive (80%) than for causal conditionals (62%; Wilcoxon test, $z = 2.67$; $p < .005$, one tail), and more predicted judgments for specific (77%) than for general (64%) conditionals (Wilcoxon test, $z = 2.24$; $p < .02$, one tail). However, the two factors interacted. As Table 3 shows, specific conditionals yielded more predicted judgments than general conditionals, but the difference was reliably larger for causal assertions (30%) than for descriptive conditionals (−5%; Wilcoxon test, $z = 2.56$, $p < .015$, two tail). This interaction probably occurred, as the participants' justifications suggested (see below), because the general causal assertions were open to counterexamples.

Two independent judges classified the justifications into four categories with a high degree of agreement ($k = 0.92$, $p < .001$), and they reconciled their discrepancies in discussion. The most frequent justifications were abductions (65%) that offered explanations for evaluations. The participants' explanations for their evaluations of the general causal conditionals were revealing. Given *If a person wins the first prize of the Euromilhões lottery, she becomes rich*, participants tended to justify their judgments of "impossible to say" by describing possibilities in which the winner loses all the money. The next most frequent justifications (21%) were paraphrases of a conditional: they reiterated its meaning, or asserted the true conditional in the case of those that were false, for example:

> I find the previously described statement [*If a dish is lasagna, then it is made with pasta*] to be true, because any type of lasagna has pasta as its basic ingredient.

Some justifications (8%) were deductive in that they mirrored the sentence completions in the justifications in Experiment 1 or were direct denials, for example:

> There is no way for Miguel to live in Lisbon and not to live in Portugal.
>
> This sentence is false because lasagna does not have rice as ingredient.

The remaining justifications (2%) were vacuous, for example:

> It's obvious.
>
> It's an unavoidable fact.

All but one of the 28 participants gave more abductive than deductive justifications, and there was one tie (Binomial test, $p = .5^{27}$, i.e., p much < 1 in a million). The difference was also reliable in an analysis by materials: abductions were more frequent than deductions for all eight sorts of conditional (Binomial test, $p = 0.5^{8}$, i.e., $p < .005$). The frequencies of the different sorts of justification were too small over the different sorts of conditional for analyses of interactions. But abductions peaked for true general causal assertions (82%) and specific descriptions whether true (89%) or false (82%). Paraphrases were fairly evenly distributed over the various sorts of conditional. And there was a tendency for deductive explanations to occur more often for judgments of falsity than for judgments of truth.

The results replicated those of Experiment 1 in that the participants tended to judge sentences as true or false a priori, albeit with performance at much less than ceiling. This decline reflects the use of the third option ("impossible to say") rather than judgments directly contrary to prediction. If one adds the percentages for the third option to the predicted category in Table 3, the resulting sums are at ceilings comparable to Experiment 1. The unified theory predicts that descriptive conditionals should be more likely to be assigned a priori truth values than causal assertions, and that specific conditionals should be more likely to be assigned a priori truth values than general conditionals. The reason is that both causal and general conditionals should be more open to counterexamples. The experiment corroborated these predictions, but it also yielded an unexpected interaction. The difference between specific and general conditionals was reliably larger for causal assertions than for descriptive assertions (see Table 3). We suspect that the interaction was a spurious consequence of the necessarily small sample of conditionals. It happened to include two true general causal assertions that participants were undecided about (50%), and two false general causal assertions that they took to be true (32%). In justifications for their judgments, which they made in their own words, all but one of the participants behaved more like scientists making abductive explanations than logicians making deductions. Because relatively few justifications referred to what was possible or what was impossible, our final experiment called for explicit judgments about them to make a direct test of the unified theory's predictions.

4. Experiment 3: Judgments and evaluations of the cases in partitions

Philosophers have distinguished between analytic truths, which depend on the meaning of assertions, and synthetic truths, which depend on knowledge of the world. And some philosophers, notably Kant (1781/1934), have argued that both sorts can be true a priori. The unified model theory also postulates that both sorts can be true a priori, and the present experiment compared them in judgments of specific descriptive conditionals. A typical trial with an analytic assertion was:

Please consider the following sentence:

If Paula has arachnophobia, then she dislikes spiders.

Do you consider that this sentence is:

True ☐ False ☐ Could be true or could be false ☐

The unified theory also postulates that compound assertions such as conditionals refer to sets of possibilities (see the Introduction). The present experiment therefore examined the patterns of participants' judgments of possibilities and impossibilities for each of the four cases in the partitions for the conditionals. A typical trial was as follows:

Please consider the following sentence in bold, and then evaluate the four given situations below as possible or impossible.

If Paula has arachnophobia, then she dislikes spiders.

Paula has arachnophobia and she dislikes spiders.	Possible ☐	Impossible ☐
Paula has arachnophobia and she does not dislike spiders.	Possible ☐	Impossible ☐
Paula does not have arachnophobia and she dislikes spiders.	Possible ☐	Impossible ☐
Paula does not have arachnophobia and she does not dislike spiders.	Possible ☐	Impossible ☐

According to the unified theory, individuals should judge an a priori true conditional, such as the one above, as possible in three cases in the partition—all those except for the second one above, which they should judge as impossible. If modulation yields a biconditional interpretation, then the third case should switch from being possible to being impossible too. For an a priori false conditional, such as:

If Paula has arachnophobia, then she likes spiders

individuals should judge as impossible only the case in the partition corresponding to the principal mental model:

Paula has arachnophobia and she likes spiders.

The other cases should be evaluated as possible. As we will explain later in the General discussion, the predicted patterns diverge from those of theories based on logic. Moreover, other a priori false conditionals, such as *If Ricardo has a daughter, then Ricardo is a mother*, should yield a different pattern of judgments:

Ricardo has a daughter and Ricardo is a mother:	Impossible.
Ricardo has a daughter and Ricardo is not a mother:	Possible.
Ricardo does not have a daughter and Ricardo is a mother:	Impossible.
Ricardo does not have a daughter and Ricardo is not a mother:	Possible.

The reason for the second judgment of impossibility is that the clause, *Ricardo is a mother*, is itself false a priori, because "Ricardo" in Portuguese refers to a male.

The experiment introduced new contents and some contingent conditionals, that is, those that should be neither true nor false a priori, and which therefore depend on the state of the world to fix their truth value, for example:

If Teresa is healthy, then she watches TV.

They should elicit judgments of possibility for all four cases in the partition.

4.1. Method

4.1.1. Participants

The participants were 146 psychology undergraduates from ISPA-IU, in Lisbon, who volunteered to take part in the experiment. They were 113 women and 33 men, with an average age of 22.5 years (*SD* = 8.23).

4.1.2. Design

The participants acted as their own controls, and for each conditional they judged its truth value (as *true*, *false*, or *could be true or could be false*) and assessed each case in its partition (as *possible* or *impossible*). The order of these two tasks was counterbalanced over the participants. The order of the four cases in the partition was also counterbalanced: The participants evaluated half of them in the order: *A & C*, *A & not-C*, *not-A & C*, *not-A & not-C*; and half of them in the reverse order. The experiment used three sorts of specific conditional: those that according to the unified theory should have an a priori truth value based on the meaning of the conditional, which we refer to as *analytic*; those that should have an a priori truth value based on knowledge of the world, which we refer to as *synthetic*; and those that were *contingent* with no a priori truth value.

4.1.3. Materials and procedure

We devised a set of 30 everyday conditionals in five categories:

Analytic truths, for example: *If Ricardo has a daughter, then he is a father.*

Analytic falsehoods: for example: *If Ricardo has a daughter, then he is a mother.*

Synthetic truths: for example: *If the gallery is in Madrid, then it is in Spain.*

Synthetic falsehoods: for example: *If the gallery is in Madrid, then it is in France.*

Contingent truth values, for example: *If Cristina goes to the beach, then she reads a book.*

As before, we devised two sets of materials and assigned participants at random to one of them to avoid any participant encountering both members of a matched pair, one true and the other false. The contingent conditionals were the same for all the participants. Each participant accordingly carried out the task with 18 sentences—six analytic (three true and three false), six synthetic (three true and three false), and six contingent. Table S3 in the Supporting Information file presents the complete materials.

The participants were tested in two separate groups. The experiment was presented in booklets, with each conditional on a different page, and in a different random order to each participant. The first page solicited the age and gender of the participant. It framed the problem, as in Experiment 2, in terms of criminology, and it also presented the key instructions for the two tasks:

> You have to judge whether sentences are true, false, or could be true or could be false, based on your knowledge.

and:

> You have to evaluate a series of cases relevant to the sentence in terms of whether they were possible or impossible.

There was no time limit to complete the experiment. After the final participant in a group had finished the experiment, the experimenter debriefed the participants in the group.

4.2. Results

The participants in the two groups to counterbalance the materials did not differ reliably in their corroboration of the predicted judgments of truth values (88% and 87%, Mann–Whitney, $z = 0.57$, $p > .5$). Likewise, the participants in the two groups to counterbalance the order of the two tasks did not differ reliably, either (87% and 84%; Mann–Whitney, $z = 0.21$, $p > .8$). Hence, we amalgamated the results for analysis. Table 4 presents the percentages of truth value judgments for the five sorts of conditional. It shows that, as in the previous studies, the participants tended to corroborate the predicted judgments of the truth values of conditionals. Only one out of the 146 participants failed to make more predicted than unpredicted judgments (Binomial test, $p <$ one in a billion). Likewise, 28 of 30 conditionals yielded more predicted than non-predicted judgments, and only one went in the opposite direction (Binomial test, $p <$ one in a million). As Table 4 shows, the synthetic conditionals (89%) yielded more predicted judgments of truth value than the analytic ones (81%; Wilcoxon test, $z = 4.35$, $p < .0001$, two tail). The main reason for this difference is that some participants thought that a person could

Table 4
The percentages of the judgments of truth values of the five sorts of conditionals in Experiment 3

		Participants' Judgments		
Predicted Judgment	Conditional Type	True	False	Impossible to Say
True	Analytic	**86**	2	12
	Synthetic	**92**	1	8
False	Analytic	9	**77**	15
	Synthetic	6	**86**	8
Indeterminate (impossible to say)	Contingent	2	11	**87**

Note. The predicted judgments are shown in bold.

be a general and a civilian, and that an orchard could be of pine trees (see sentences 2′ in Set I and Set II of Table S4 in the Supporting Information file).

Table 5 presents the percentages of the most frequent patterns of evaluations of the four cases in the partition as possible or impossible. The predicted patterns occurred more often for the synthetic conditionals (62%) than for the analytic conditionals (54%; Wilcoxon test, $z = 4.09$, $p < .001$). The difference appears to reflect merely the unexpected evaluations of the assertions about the general and the orchard. Overall, the predicted patterns shown in the table occurred more often than chance. There are 16 possible patterns of evaluation of the four cases in the partition for a conditional, and all 16 of them occurred at least once in the experiment. Hence, the probability that by chance an evaluation of the four cases matches the unified theory's predictions is 1/16. Out of the 146 participants, 88 made more predicted patterns of evaluation than non-predicted patterns, 53 made more non-predicted than predicted patterns of evaluation, and there were 5 ties (Binomial test with a prior probability of 1/16 yields p close to 0). Participants accordingly differed one from another in their propensity to evaluate the possibilities for conditionals, but we do not know what underlying individual difference might be responsible for this result. There were 28 conditionals out of 30 that yielded more predicted evaluations than not, and the two conditionals that, as we mentioned earlier, yielded more unpredicted evaluations than not (Binomial test, $p < 1$ in a billion). The pattern in Table 5 in which only two cases are evaluated as possible (*A and B*, and *not-A and not-B*) is not contrary to the unified theory, because it corresponds to a biconditional interpretation. We chose not to treat it as corroborating the theory to make the test more stringent. The conditional *If Ricardo has a daughter then he is a mother* has a *then*-clause that should be false a priori, but it should yield a different pattern of evaluations for such conditionals, which we described earlier. The general patterns of evaluation, however, corroborated the unified theory.

Table 5
Percentages of the more frequent patterns of evaluations made for conditionals predicted to be judged true, false, and contingent in Experiment 3

The Predominant Patterns of Cases in the Partition Judged to Be Possible					
Predicted Sort of Conditional	A B ¬ A B ¬ A ¬ B	A ¬ B ¬ A B ¬ A ¬ B	A B A ¬ B ¬ A B ¬ A ¬ B	A B A B ¬A ¬ B	
Judged to be true					Sum
Analytic	**63**	0	6	22	92
Synthetic	**75**	0	6	10	91
Judged to be false					
Analytic	16	**35**	7	11	69
Synthetic	18	**48**	6	8	81
Judged to be contingent	18	0	**62**	12	92

Note. Predicted evaluations are in bold, and cases such as: *not-A & B*, are abbreviated in the table as: ¬ *A B*.

5. General discussion

Philosophers have long argued that certain assertions are true a priori. They took the truth of an assertion such as:

No bachelor is married

to follow directly from the meanings of the words it contains. It is an *analytic* truth. This view became much less popular among philosophers after the publication of Quine's seminal essay, "Two dogmas of empiricism" (see Quine, 1953). He allowed that certain sentences expressed a priori truths—those that were true in virtue of logic alone, such as:

No unmarried man is married.

But he argued that analytic truths depending on meaning, such as the preceding example, are indistinguishable from contingent assertions. Readers might suppose that one could replace "unmarried man" in the preceding sentence with the synonymous term, "bachelor," and so infer that the first sentence above is true, too. But to do so is rely on the concept of synonymy, which for Quine was just as problematic as analyticity. In consequence, the distinction between assertions that are true a priori in virtue of their meaning and those that depend on matters of fact is "an unempirical dogma of empiricism" (Quine, 1953, p. 37). Our experiments have shown collectively that well over 600 individuals, naive about philosophical niceties, judged that some conditional assertions are true a priori, and that some are false a priori. That is, their truth values do not depend on factual evidence. The percentages of such predicted judgments varied over our three experiments in a way that makes sense. In Experiment 1, they were at ceiling around 98%, but the participants could evaluate the conditionals only as either true or else false. In Experiment 2, they fell to around 71%, but the participants now had the option to respond, "impossible to say," and the experiment included both causal and general conditionals for which, as the theory predicts, it is easier to think of counterexamples than for their counterparts, descriptive and specific conditionals. In Experiment 3, the predicted judgments were around 85%; it too included a third option, "could be true, could be false," but made no use of causal or general assertions—all the conditionals were specific and descriptive, and so less open to counterexamples. Hence, whatever else the idea of a priori truth values might be, it is no longer an unempirical dogma—experimental findings support it.

Of course, not all of the participants' judgments fitted the predictions. For example, nearly a third of the participants failed to judge as false the assertion:

If João is in Azores, then he is in on the mainland.

Why did such errors occur? Someone once asked Dr. Samuel Johnson why he had erroneously defined "pastern" in his famous dictionary as the knee of a horse. (In fact, it is the part that slopes down from the fetlock to the hoof.) He replied, "Ignorance, Madam, pure ignorance." The same reason, we suspect, lies behind many of our participants' apparent errors—that is, our ignorance about theirs.

Our results elucidate the meanings of conditional assertions. According to the unified model theory, sentential connectives refer to a conjunction of possibilities and impossibilities. An unmodulated conditional of the sort, *If A, then C*, refers to three cases as possible:

A	&	C
Not-A	&	C
Not-A	&	not-C

and one case as impossible:

A & not-C.

The content of some conditionals modulates their interpretation so that they are treated as biconditionals, equivalent to *If, and only if, A then C*, for example:

If Emma is illiterate, then she is unable to read.

Modulation can yield a variety of other interpretations, which we discussed in the Introduction. But a basic principle is that if a conditional refers to more than one case in the partition as possible, one of these cases must correspond to the one explicit mental model of the conditional in its two mental models (Johnson-Laird & Byrne, 2002, p. 660):

A C
. . .

In simple tasks, individuals are able to flesh out these mental models of possibilities into fully explicit models, which we show in the order in which they tend to enumerate them (see, e.g., Barrouillet et al., 2000; Johnson-Laird & Byrne, 2002):

A	C
Not-A	not-C
Not-A	C

The third of these cases is impossible for a biconditional interpretation. It follows from this account that when individuals justify their judgments of the a priori truth values of conditionals and biconditionals, they should complete a sentence of the sort:

It is impossible that A and __

by asserting *not-C* for conditionals that they judge to be true, and by asserting *C* for those that they judge to be false. Nearly all of them did so (around 96% in Experiment 1). It also follows that they should be likely to make patterns of judgments of conditionals that sometimes reflect a biconditional interpretation, and they do, as Table 5 shows for Experiment 3.

When participants made their own justifications for their judgments of a priori truth values, they acted more like scientists than logicians. They abduced explanations for the truth and for the falsity of conditionals (Experiment 2). The bias toward abduction is congruent with a general tendency in human reasoning: Individuals are highly adept in formulating explanations and can do so with a facility well beyond any current computer model. Indeed, in a task that allowed either an abductive or deductive response, they were biased toward abduction (Johnson-Laird, Girotto, & Legrenzi, 2004).

When participants in Experiment 3 had to make an explicit judgment of the four cases in the partition for a conditional, a striking relation occurred between these judgments and those of whether a sentence is true a priori or false a priori. For a conditional that is judged to be true, such as:

If Rui is obese, then he is overweight

the typical pattern of evaluations corresponded to the unified theory's predictions. Individuals judged that each case is possible except for:

Rui is obese and he is not overweight

which they evaluated as impossible. But some individuals made a biconditional interpretation, yielding a further impossibility:

Rui is not obese and he is overweight.

For a conditional that tends to be judged as false a priori, for example:

If Paula has arachnophobia, then she likes spiders.

the typical pattern of evaluation contains only one case that is impossible, the one corresponding to *A & C* (the principal mental model of the conditional):

Paula has arachnophobia and she likes spiders.

The other three cases are all possible. Some conditionals, however, are false a priori because their *then*-clauses are false a priori, for example:

If Ricardo has a daughter, then he is a mother.

They yielded the predicted pattern of evaluation in which a further case was evaluated as impossible:

Ricardo does not have a daughter and he is a mother.

And, although we did not include any such conditionals in Experiment 3, other conditionals should be judged false a priori, because their *if*-clauses are false, for example:

If Ricardo is a mother, then he has a child.

They too should yield a different pattern of evaluation in which both cases in which Ricardo is a mother are evaluated as impossible. The unified theory's predictions are summarized in Table 6 so that we can compare them with accounts based on logic.

Psychological theories based on logic are major alternatives to the unified theory (see, e.g., Braine & O'Brien, 1998; Rips, 1994). These accounts as they are currently formulated cannot cope with the concepts of possibility and impossibility. But certain modal logics do combine these concepts with sentential logic (see Hughes & Cresswell, 1996; Ch. 2), and a pioneering psychological theory incorporated principles of such a modal logic (Osherson, 1976). As we pointed out in the Introduction, the logical analog of a conditional, *If A, then C*, is material implication: *A materially implies C*. It is true provided that any case in the partition other than *A & not-C* is true; likewise, it is false in the complement of cases: *A & not-C* is true, but the other three cases in the partition are false (see Jeffrey, 1981). The patterns of evaluations for true conditionals are compatible with logic, but they diverge from it for false conditionals. Table 6 summarizes the differences. In logic, if a proposition is true, then it is possible—an assumption that is implausible in daily life, because individuals are unlikely to argue: *It is raining*; therefore, *it is possible that it is raining* (see Karttunen, 1972). Nevertheless, granted that truth implies possibility, then it follows in logic that if a proposition is impossible, then it is false. Our results therefore count against theories based on logic. Wherever a conditional elicits an "impossible" judgment, its truth value in logic should be false. As Table 6 shows, there are two cases to the contrary, for the *A and not-C* and *not-A and C* cases, respectively. Likewise, people judge that a conditional such as *If Mary has the flu, then she is healthy* is false a priori. It follows in logic that the case in which *Mary does not have the flu* is false (see Table 6). So, if you judge a conditional as false, you are thereby asserting the truth of its *if*-clause. Hence, in logic, there is a simple proof that God exists. The conditional: *If God exists, then atheism is correct*, is false, and so it thereby that its *if*-clause is true: *God exists*. The inference is merely the mirror image of the "paradox" of material implication that we described in the Introduction. In contrast, the falsity of a conditional in the unified theory has no such implication: The cases in which its *if*-clause is negated are merely possible (see Table 6).

The other main alternative to the unified model theory is probabilistic logic, or p-logic for short (see, e.g., Adams, 1998; and for a defense of its psychological reality, see, e.g.,

Table 6
The analysis of the falsity of material implication in logic and the falsity of conditionals in the unified model theory

		Three examples of the falsity of *If A, then C* in the unified theory		
The four conjunctive cases in the partition	The truth values for the falsity of *A materially implies C* in logic	If Mary has the flu, then she is healthy	If Ricardo has a daughter, then he is a mother	If Ricardo is a mother, then he has a child
A C	False	Impossible	Impossible	Impossible
A not-C	True	Possible	Possible	Impossible
not-A C	False	Possible	Impossible	Possible
not-A not-C	False	Possible	Possible	Possible

Cruz et al., 2015; Evans, 2012; Pfeifer & Kleiter, 2009; and for a critique, see Johnson-Laird et al., 2015). Its central equation is that the probability of a conditional, *If A, then C*, equals the conditional probability of *C* given *A,* granted that *A* has a non-zero probability. If individuals judge that a conditional is true, it follows that the probability of *C* given *A* equals unity. This value in turn determines that *A & C* is possible, and that *A & not-C* is impossible. This pattern is compatible with the unified theory. The equation has no implications, however, for whether cases of *not-A* are possible or impossible. Likewise, the status of these cases is still more obscure for conditionals that are false. Proponents of the probabilistic approach often advocate a suppositional theory in which individuals assess a conditional by adding its *if*-clause hypothetically to their stock of beliefs and then evaluating its *then*-clause. As a consequence, conditionals do not have a truth value when their *if*-clauses are false (Evans, 2007, p. 56). This view runs into many difficulties (see, e.g., Johnson-Laird et al., 2015). The most salient one in our results is that individuals judge such cases to be possible for true conditionals in Experiment 3 (see Table 5). But, if a case is possible, then it could be true—an eventuality contrary to the suppositional theory.

The unified theory treats the concept of possibility as basic: a person can judge that a proposition is possible without having to consider its probability, but the theory does explain the origins of the numbers in the probabilities that numerate individuals assign to assertions (see, e.g., Khemlani, Lotstein, & Johnson-Laird, 2015). It seems that theories based on p-logic treat probabilities as basic, and so the status that they accord to possibilities is unclear. Likewise, they offer no account of how numbers come to be assigned to probabilities.

Could it be that Leibiz, Kant, Carnap, and our participants are mistaken in their judgments, which are akin to medieval dogmas, such as that the Earth is flat? The controversy about the Earth was settled by observation and inference. In our view, so too is the controversy about a priori truth values. No court of appeal exists higher than one that asks individuals to judge the truth or falsity of assertions. Given a conditional, such as:

If Mary has the flu, then she is healthy

most people are likely to respond, as did our participants, that the claim is false, and that it is impossible for Mary to have the flu and to be healthy. People know that flu is an illness, and that a person who is ill is not healthy. And that is why they judge that it is impossible that Mary has the flu and is healthy. Skeptics, like Quine (1953), may argue that no assertion is immune to revision. But this one seems quite incontrovertible. Flu is an illness, and one that can kill. Whatever changes it may undergo, which in turn may change what "flu" refers to, are irrelevant. Judgments in the marketplace or in a court of law are based on current knowledge and the current meanings of words. And our participants judged the conditional to be false. Indeed, some conditional assertions are true in virtue of knowledge, some are false, and some could be true or false depending on the evidence. Our studies have concerned indicative conditionals, but a priori truth values should occur for other compound assertions, such as disjunctions, and even for counterfactual conditionals such as:

If Newton had been right, then mass would have had no effect on light.

The justifications that individuals adduce for their judgments tend to explain them rather than to concern possibilities. Yet, when they evaluate each case in the partition as possible or impossible, the majority of the resulting patterns of evaluation bear out the unified model theory. It implies that at the heart of the meaning of conditionals are conjunctions of possibilities and impossibilities.

Acknowledgments

We thank the Portuguese Foundation for Science and Technology (FCT) for the grant to the William James Center for Research (UID.PSI.04810/2013) to the first author. We thank CTT-Correios de Portugal (Portuguese Post), for helping us by distributing the link for Experiment 1 to their employees. We thank Marta Couto, João Marques, and Pedro Fernandes, for their work in helping to develop the materials, to carry out the experiments, and to analyze the results. We are also grateful to the following colleagues for advice: Ruth Byrne, Geoff Goodwin, Sangeet Khemlani, and Juan García-Madruga.

References

Adams, E.W. (1998). *A primer of probability logic*. Stanford, CA: Center for the Study of Language and Information.

Barrouillet, P., Grosset, N., & Lecas, J.-F. (2000). Conditional reasoning by mental models: Chronometric and developmental evidence. *Cognition*, *75*, 237–266. doi:10.1016/s0010-0277(00)00066-4

Braine, M. D. S., & O'Brien, D. P. (Eds.) (1998). *Mental logic*. Mahwah, NJ: Erlbaum.

Carnap, R. (1947). *Meaning and necessity: A study in semantics and modal logic*. Chicago: University of Chicago Press.

Chomsky, N. (1977). *Essays on form and interpretation*. New York: North-Holland.

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in psychology*, *6*, 109. doi:10.3389/fpsyg.2015.00192

Evans, J.St.B.T. (2007). *Hypothetical thinking*. Hove, England: Psychology Press.

Evans, J.St.B.T. (2012). Questions and challenges for the new psychology of reasoning. *Thinking & Reasoning*, *18*, 5–31.

Evans, J.St.B.T., & Over, D.E. (2004). *If*. Oxford, England: Oxford University Press.

Goodwin, G., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, *112*, 468–493. doi:10.1037/0033-295x.112.2.468

Hinterecker, T., Knauff, M., & Johnson-Laird, P. N. (2016). Modality, probability, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(10), 1606–1620. doi:10.1037/xlm0000255

Hughes, G. E., & Cresswell, M. J. (1996). *A new introduction to modal logic*. London: Routledge.

Hume, D. (1978). A treatise of human nature. In P. H. Nidditch (Ed.), second edition. Oxford, England: Oxford University Press. (Originally published, 1739.)

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

Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, England: Cambridge University Press. Cambridge, MA: Harvard University Press.

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

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

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

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review*, *111*, 640–661. doi:10.1037/0033-295x.111.3.640

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, *19*, 201–214. doi:10.1016/j.tics.2015.02.006

Juhos, C., Quelhas, C., & Johnson-Laird, P. N. (2012). Temporal and spatial relations in sentential reasoning. *Cognition*, *122*, 393–404. doi:10.1016/j.cognition.2011.11.007

Kant, I. (1934). *Critique of pure reason.* Trans. J. M. D. Meiklejohn. New York: Dutton. (Originally published 1781.)

Karttunen, L. (1972). Possible and must. In J. P. Kimball (Ed.), *Syntax and semantics*, Vol. 1. (pp. 1–20.) New York: Seminar Press.

Khemlani, S., & Johnson-Laird, P. N. (2013). The processes of inference. *Argument and Computation*, *4*, 4–20. doi:10.1080/19462166.2012.674060

Khemlani, S., Lotstein, M., & Johnson-Laird, P. N. (2015). Naive probability: Model-based estimates of unique events. *Cognitive Science*, *39*, 1216–1258. doi:10.1111/cogs.12193

Leibniz, G.W. (1686) *Discourse on Metaphysics and the Monadology* (trans. George R. Montgomery). Amherst, NY: Prometheus Books, 1992. (Originally published 1686.)

Miller, G.A., & Johnson-Laird, P.N. (1976). *Language and perception*. Cambridge, England: Cambridge University Press. Cambridge, MA.: Harvard University Press.

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

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, England: Oxford University Press.

Orenes, I., & Johnson-Laird, P. N. (2012). Logic, models, and paradoxical inferences. *Mind & Language*, *27*, 357–377. doi:10.1111/j.1468-0017.2012.01448.x

Osherson, D.N. (1976). *Logical abilities in children, Vol. 4: Reasoning and concepts*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic*, *7*, 206–217.

Quelhas, A. C., & Johnson-Laird, P. N. (2017). The modulation of disjunctive assertions. *Quarterly Journal of Experimental Psychology*, *70*(4) 703–715, doi:10.1080/17470218.2016.1154079

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. doi:10.1080/17470210903536902

Quine, W.V.O. (1953). Two dogmas of empiricism. In W. V. O Quine (Ed.), *From a logical point of view* (pp. 20–46). Cambridge, MA. Harvard University Press. (Originally published, 1951).

Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.

Siegel, S., & Castellan, N. J., Jr (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw Hill.

Steinberg, D. D. (1970). Analyticity, amphigory, and the semantic interpretation of sentences. *Journal of Verbal Learning and Verbal Behavior*, *9*, 37–51. doi:10.1016/s0022-5371(70)80006-8

Steinberg, D. D. (1972). Truth, amphigory, and the semantic interpretation of sentences. *Journal of Experimental Psychology*, *93*, 217–218. doi:10.1037/h0032510

Steinberg, D. D. (1975). Semantic universals in sentence processing and interpretation: A study of Chinese, Finnish, Japanese, and Slovenian speakers. *Journal of Psycholinguistic Research*, *4*, 169–193. doi:10.1007/bf01066925

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. The 48 conditional sentences used in Experiment 1 (in English and in the original Portuguese).

Table S2. The 16 conditional sentences used in Experiment 2 (in English and in the original Portuguese).

Table S3. The 30 conditional sentences used in Experiment 3 (in English and in the original Portuguese) and the percentages of their predicted evaluations in the experiment.